LUC BISSONNETTE

ESSAYS ON SUBJECTIVE EXPECTATIONS AND STATED PREFERENCES

ESSAYS ON SUBJECTIVE EXPECTATIONS AND STATED PREFERENCES

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in zaal AZ17 van de Universiteit op maandag 23 januari 2012 om 16.15 uur door

Luc Martin Bissonnette

geboren op 4 december 1981 te Loretteville, QC, Canada.

Promotor
Prof. dr. Arthur van Soest

Promotiecommissie
Prof. dr. Rob Alessie
Dr. Charles Bellemare
Dr. Pierre-Carl Michaud
Dr. Martin Salm

To Marie-France, the love of my life.

To my parents, without whom I would not be here.

And to all those who still believed in me when I, myself, could not believe anymore.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

Chapter 2
Bellemare, Bissonnette, and Kröger (forthcoming): "Flexible Approximation of Subjective Expectations Using Probability Questions," *Journal of Business and Economic Statistics*.

Chapter 3
Bellemare, Bissonnette, and Kröger (2010): " Bounding Preference Parameters Under Different Assumptions About Beliefs: A Partial Identification Approach," *Experimental Economics*, 13 (3), 334–345.

Chapter 4
Bissonnette and van Soest (2011): "The Future of Retirement and the Pension System: How the Public's Expectations Vary over Time and across Socio-Economic Groups," *CentER Discussion Paper Series*, (2011-065).

AVANT-PROPOS

Despite the lack of natural ability, I did have the one element necessary to all early creativity: naïveté, this fabulous quality that keeps you from knowing just how unsuited you are for what you are about to do.

— Steve Martin, *Born Standing Up*

Time sure flies fast! While most of my friends were busy founding (and funding) families, I was busy writing this blue booklet. And now, after 5 years that felt like 15, it is finally time to write this *avant-propos* (French for *foreword*, you are welcome). This section is about the personal part of this saga, so let's write it at the image of its creator: entertaining (at least, that's the intent), falsely modest, and lacking the deferential tone proper to academia. And grateful. Mostly grateful.

Very grateful, first and foremost, to my supervisor Arthur. What to say about Arthur? I could talk about his astonishing intelligence and his passion for empirical research, but I would not know what to say that has not been said a thousand times... And I believe that his greatest quality lies elsewhere: Arthur is a great mentor. I am always amazed at how he manages to help his students achieve the best of their potential, no matter how (un)skilled we are. I will remember for the rest of my life the meetings we had together, where he would nod his (dis)approval in a simple manner. I must admit that it took me some time to adapt to working with Arthur. I first had to get used to some Dutch subtleties lost in their English translations (the real English, not the American one). I had to learn that "it is fine" means "it really needs more work" and that "it is good" means "it is great". Pretty much everything I know in econometrics was taught to me with his pedagogical catch-phrase: "I am not sure about this", which can roughly be translated as "see, you are wrong here".

I am also grateful to those who first fuelled my zealous love for research: Charles and Sabine. They believed in my skills enough to help me start a real research career. I will always be indebted to them for giving me the opportunity to study in a good PhD programme – or as some call it in Tilburg, a *top* programme.

I also want to thank the members of my doctoral committee who provided a lot of insightful comments. Two members of the committee did not co-author this thesis, meaning that they got to read the whole thing without any direct benefits. After reading all the chapters of this dissertation and our very thick *Netspar Panel Paper #18*, Rob officially became my "main readership", having read pretty much everything that I ever wrote in my short academic life. Martin, who was also on the committee for my formerly-known-as-m-phil-thesis, is a close

second place. Martin was also of great moral support while I was working on my thesis, but he got a nice cookbook and some advices on wines in exchange, so I'd say we are even for that part. The last member of my committee, Pierre-Carl, deserves special thanks namely for inviting me to spend a few much-needed months under the sun in Southern California and for a few drinks at Rick's. I also take this opportunity to thank Michael Hurd and the team of researchers who welcomed me at RAND for their hospitality.

Long after my memory will erase the taste of croquettes or forget the Dutch obsession with cheap satay sauce, I hope that I will remember all the kind people who made my academic life a little less depressing. I got to live with a lot of officemates, and I want to thank Fangfang's epic zeal, Cristina's infinite kindness, Juan Juan's (or was it Ms. Kai?) shy cleverness, Hanka's sweet "unappropriateness" (and *questionable* driving skills), Salvatore's elusive presence and WeoJong Kim. Beyond the boundaries of my various offices, I cannot stay silent about David's sharp sarcasm, Jan's unshakable enthusiasm, Marco's under-appreciated genius, Rasa's timid intellect, Guillaume's quiet wisdom, Ting's lively philosophy, Jarda's energetic *joie de vivre* (and unhealthy TV addiction), Peter's fast-paced English (even compared to mine), and Ying's cheerful spirit. Some special thanks to Lisanne, Mohammed, Patrick, Nathanael, Peter and all the "Welcoming Dutch" who helped me understand this peculiar flat land in which I landed. Many young professors in Tilburg were of great support to me, and I hope that Tobias, Otilia, and Meltem realize how much I valued their help and support. I am particularly indebted to Katie, who took a lot of her precious time to listen to my depressed whining and to guide me in the fog of academic life.

Finally, this thesis would not exist without the support of my wonderful life-partner Marie-France. She made even more sacrifices than I did over these four years when she decided to come with me to live in Tilburg. As I write these lines, we just passed the milestone of ten happy years together. I clearly got the upper hand of this deal. My parents, Gilles and Sylvie, were also very supportive (both morally and financially) through my 25-year career as a student. I am very grateful for that too.

Some people say I leave Tilburg as an inconsiderate prodigal child. I do not feel that way, and I am very grateful to those who helped me become a better researcher. Besides I believe that a small part of me will always stay in Tilburg. A small part of me I used to call "Sanity".

Thank you for reading me.

Luc Bissonnette
Québec, December 10, 2011

CONTENTS

1 INTRODUCTION

We all know that smoking increases mortality risks, but how dangerous is smoking for our health? We also know that the world population is aging, but what is the likely impact of this on our retirement plans? What are the probabilities that we will be the victims of car theft or robbery? Questions about uncertain consequences are important when we decide whether to smoke or not, how much to save, or whether or not to buy insurance. However, few of us are able to provide accurate answers. As a matter of fact, every day we make decisions under uncertainty without a perfect knowledge of the risks we face. Each of us may have a different perception of the probability of the realization of future events. These heterogeneous predictions are known to economists as *subjective expectations*. The study of this heterogeneity in expectations is the unifying theme of the chapters of this thesis.

There are many factors explaining this heterogeneity. On the one hand, we live in a complex reality and it is hard to accurately assess all the probabilities needed to take important decisions. Agents will typically not know, for example, the precise probability they will be in a car accident or will have cancer, but will still decide to insure themselves or not. On the other hand, they have a lot of private information concerning their own driving habits or medical history, and this information offers them the possibility to form more accurate expectations than an outside observer without this knowledge. Even if agents form objectively valid expectations, they may very well have heterogenous expectations.

Despite the fairly uncontroversial view that there is heterogeneity in expectations, the standard assumption in economics is that economic agents are able to form what has been known as *rational expectations* since the work by Muth (1961). The rational expectations assumption implies that while economic agents cannot perfectly forecast the future, they know the objectively valid distribution of uncertain events. This assumption became a prominent part of the economic literature, particularly with the work of macroeconomists like Lucas (see Pesaran, 1987, for a discussion). Skepticism concerning rational expectations is not a new concern in economics. An early discussion of the matter can be found in an essay by the mathematician Ramsey (1926) entitled *Truth and Probability*.[1] In his essay, Ramsey wanted to contrast his personal view that people were not able to form accurate beliefs with the

1 The essay, written in 1926, was published posthumously in 1931 as a chapter in *The Foundations of Mathematics and other Logical Essays*.

view held by Keynes that they could form objectively valid beliefs. He wrote:

> Mr Keynes[2] starts from the supposition that we make probable inferences for which we claim objective validity; we proceed from full belief in one proposition to partial belief in another, and we claim that this procedure is objectively right, so that if another man in similar circumstances entertained a different degree of belief, he would be wrong in doing so. [...]
>
> [A fundamental criticism of Mr Keynes' views] is the obvious one that there really do not seem to be any such things as the probability relations he describes. He supposes that, at any rate in certain cases, they can be perceived; but speaking for myself I feel confident that this is not true. I do not perceive them, and if I am to be persuaded that they exist it must be by argument... [pp. 160-161]

Ramsey's view that there is no such thing as an intrinsic process linking precisely objective realities and subjective probabilities may sound trivial for those who never studied economics. I suspect most individuals to feel like him: if I cannot form rational beliefs, why would others be able to? Why assume that expectations are rational when we have compelling evidence that individuals do not have such accurate beliefs?

One of the reasons for relying on the rational expectations assumption in empirical analysis is that data on expectations are rarely available.[3] Beliefs are never directly observed and can usually not be inferred by simply observing the agents' choices. Because many combinations of preferences and beliefs may lead to the same decision, it is impossible to disentangle the effects of these two components on the decision process. This point was illustrated by Manski (2002), who studied the identification of a model of decision making in a simple experimental game of proposal and response. If it is impossible to disentangle expectations and preferences, assumptions on expectations are needed in order to estimate preference parameters. In most cases, the rational expectations assumption is seen as a plausible assumption, or at least as the most plausible one. In some cases, parametric assumptions can allow for heterogeneity in beliefs (e.g. Rust and Phelan, 1997), but as pointed out by Bernheim (1989), however, identification in such cases relies solely on the functional form. An approach with weaker assumptions that would allow direct identification of the role

2 J.M. Keynes, A Treatise on Probability (1921), footnote from the original citation.

3 Another simple reason that comes to mind is that the rational expectation hypothesis simplifies many problems and leads to elegant analytical solutions in theory. Given that the core focus of this thesis is empirical, I will omit this aspect from the discussion.

of beliefs in a decision requires additional information on subjective expectations. However, beliefs are rarely observed.

To solve the problem that expectations are usually unobserved, a recent strand in the empirical micro-economic literature aims at measuring expectations directly using survey questions (see Manski, 2004, for a review of this topic). While expectations data have been studied for some time (e.g. Hall and Johnson, 1980; Hamermesh, 1985), elicitation of expectation data became an increasingly important field in economics since the early 1990s, mostly due to the contribution of Manski (e.g. 1990; 1993). In these studies, survey respondents are asked explicitly to report the probability that a certain outcome will realize. This information on individual beliefs allows the role of expectations in an economic model to be isolated, which in turn makes it possible to identify preferences parameters in an econometric estimation. My thesis contains papers that were built on this precept: that it is possible to study empirically subjective expectations based on subjective data.

As I write these lines, papers on subjective expectations are a flourishing part of the economic literature. Reviewing all the work done in this field is clearly beyond the scope of this Introduction. There are now more and more surveys eliciting information on subjective expectations and some of these important surveys, such as the HRS, have included probability questions on this type for almost two decades (see Hurd, 2009, for a review concerning the HRS). In some cases, the realizations corresponding to these forward-looking questions are now known, and it is possible to discuss whether or not expectations were accurate. While the question of whether or not expectations can be characterized as rational is still part of the new literature (e.g. Benítez-Silva, Dwyer, Gayle, and Muench, 2008), the idea that agents do not necessarily form objectively valid expectations is now widely accepted, which has led to many additional research topics. Researchers are interested to know if expectations, beyond their rationality, have predictive power for the observed outcome (e.g. Hurd and McGarry, 2002) or whether subjective expectations help to predict behavior that should be affected by these beliefs (e.g. Nyarko and Schotter, 2002). Some papers focus on practical issues such as how to elicit subjective expectations (e.g. Delavande and Rohwedder, 2008), others develop econometric methods to analyze elicited subjective expectations (e.g. Dominitz and Manski, 1996) or to take into account measurement errors in reporting and rounding of answers (e.g. Manski and Molinari, 2010; Kleinjans and van Soest, 2010).

This dissertation contains five core chapters. The research questions are in line with those mentioned above. In some chapters, we are interested in the study of expectations per se and study, for instance, who are the individuals expressing higher probabilities that a given outcome will realize. Many of the chapters discuss important econometric issues in the analysis of subjective expectations, proposing ways to estimate models with subjective expectations or discussing the importance of subjective expectations to identify key preference parameters of a model. In some cases, the implication of heterogeneity in subjective beliefs for relevant economic decisions will be discussed. One chapter proposes a test of rational expectations. Finally, one chapter discusses an approach to estimate a model without having to elicit various subjective expectations that would usually cause problems.

The thesis is divided into two parts, discussed below in more detail. The first part, grouping two of the five chapters, contains work where the main focus is the analysis of subjective expectations from the point of view of an applied econometrician. In the second part, the focus is on the use of subjective expectations and subjective data in the field of the economics of aging.

Part I: Essays on the econometrics of subjective expectations

The first part of this thesis includes the two chapters written jointly with my supervisors during my years as a masters student in Québec City, Charles Bellemare and Sabine Kröger. In essence, these chapters are concerned with methodological issues concerning the analysis of subjective expectations, more than with the study of expectations per se.

Chapter 2 introduces a flexible method to estimate a subjective probability distribution function when many points along this distribution are observed. Prior to this reasearch, the most common way to estimate these distributions, proposed by Dominitz and Manski (1996), was to assume a parametric distribution and to find the parameters leading to the distribution with the best fit to the available points. We propose a more flexible, nonparametric approach relying on weaker assumptions: an intrapolation with cubic splines. We show that this method is robust to measurement errors due to rounding, performs almost as well as efficient non-linear least squares estimation when the correct parametric distribution is known, and outperforms non-linear least squares estimation when the wrong parametric assumption is chosen.

Chapter 3 is a discussion of the identification power of various assumptions on beliefs in a simple model of decision making in an experimental game. Experimental economics, by the simplicity of the decisions involved, is an interesting setting to study the role of subjective expectations in decision making (e.g. Nyarko and Schotter, 2002; Bellemare, Kröger, and van Soest, 2008).The decision of interest is the behavior of a participant who has to decide whether to invest or not in a simple *investment game* inspired by the protocol of Berg, Dickhaut, and McCabe (1995). We compare the identification of the model under various assumptions, and show that unless we observe the heterogeneous subjective beliefs, we can only have partial identification – even under the assumption of rational expectations.

Part II: Subjective beliefs and economics of aging

The second part of the thesis, grouping Chapters 4, 5, and 6, focuses on the analysis of subjective expectations in the economics of aging. The economics of aging is one of the fields where the study of subjective expectations represents an important part of current research (see Hurd, 2009; Bissonnette and van Soest, 2010, for reviews on this topic). This is not surprising since, because of the inter-temporal nature of the many topics studied in this discipline, beliefs play a key role. For instance, an individual's expectations concerning his retirement age or his life-expectancy are important factors in decisions to save for retirement. Moreover, as may countries are moving away from defined benefits plans to defined contribution plans, the burden of decisions concerning retirement plans requires a level of financial knowledge that is neither trivial nor, in many cases, part of the standard curriculum at school. Research on financial literacy shows that many respondents do not understand this complex economic world (see for instance Lusardi and Mitchell, 2007). This lack of financial knowledge may lead to erroneous expectations, which may in turn lead to suboptimal saving decisions.

The idea that respondents may have heterogeneous beliefs concerning their pensions is the topic studied in Chapter 4. This chapter is based on joint work with Arthur van Soest. In this chapter we study how expectations concerning changes in social security vary over time and among socio-economic and demographic groups. We show that despite the fact that many respondents adapt their beliefs concerning an upcoming reform of the Dutch public pension system, some individuals in the population put themselves at risk, not being able to forecast coming changes.

Expectations of life-expectancy is the central theme of Chapter 5, written jointly with Michael Hurd and Pierre-Carl Michaud. In this chapter, we try to characterize the full subjective survival curves (i.e. the probability of survival as a function of age) of survey respondents based on their self-reported survival at a target age. This paper is an extension of the model introduced by Gan, Hurd, and McFadden (2005). Our main contribution is to develop a procedure relying on within-sample mortality rather than on nationally representative life-tables. Doing so allows us to avoid a selection problem that occurs if survey respondents do not have nationally representative mortality risks and allows us to include a socio-economic gradient even in cases where life-tables are not published (e.g. assessing mortality by education level). As I mentioned at the beginning of this chapter, part of the private information that we have may lead to the formation of beliefs that are rational at the individual level, despite being different from what we can observe at the aggregated level. Some individuals have better objective survival probabilities than others. It has already been shown that the subjective survival probabilities were predictive of one's survival status (Hurd and McGarry, 2002). In this chapter, we develop a way to test whether subgroups of the population are rational. Our approach is to compare objective mortality risk within our sample with self-reported survival probabilities. We find a general optimism in the population, mostly driven by a group of very optimistic respondents not matched by a group of equally pessimistic respondents.

The last chapter about stated-preferences, Chapter 6, is rather different in nature from the others. This chapter uses stated-preference data to estimate a simple model of retirement and savings. Respondents were asked to evaluate a series of retirement scenarios, stating their preferences concerning retirement plans. These simple scenarios were written in a way that ruled out most of the uncertainty faced by the respondents. As a matter of fact, this chapter took its root precisely from the fact that there is no need to characterize expectations held by the respondents, allowing for estimation of preferences without having to be concerned that decisions may be influenced by other factors. Therefore, I end the thesis with an alternative solution to the estimation of models where heterogeneity in expectations, rational or not, may lead to biased estimates of preference parameters. The main objective of the paper is to assess the implications of the hypothetical scenarios in terms of a life-cycle model. To illustrate the results, I simulate a delay of two years in eligibility to social security and predict the variation in current savings and delay in retirement for the survey respondents. The main conclusion from this chapter is that the use of stated preferences leads to plausible estimates of all parameters of interest. Using these data, I find a discount factor between 0.95 and 0.97 and a coefficient of relative risk aversion around 1.15, in line with estimates from Hurd (1989). It would be interesting, in future work, to validate that respondents' stated preferences are correlated with their actual behavior and to assess if combining stated preferences with more traditional revealed preferences can help in estimating a more traditional retirement model.

BIBLIOGRAPHY

Bellemare, C., S. Kröger, and A. van Soest (2008): "Measuring Inequity Aversion in a Heterogeneous Population using Experimental Decisions and Subjective Probabilities," *Econometrica.*, 76(4), 815–839.

Benítez-Silva, H., D. S. Dwyer, W.-R. Gayle, and T. J. Muench (2008): "Expectations in Micro Data: Rationality Revisited," *Empirical Economics*, 34(2), 381–416.

Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, Reciprocity and Social History," *Games and Economic Behavior*, 10, 122–142.

Bernheim, B. D. (1989): "The Timing of Retirement: A Comparison of Expectations and Realizations," in *NBER Chapters*, pp. 335–358. National Bureau of Economic Research, Inc.

Bissonnette, L., and A. van Soest (2010): "Retirement Expectations, Preferences, and Decisions," *Netspar Panel Paper*, 18, 1–98.

Delavande, A., and S. Rohwedder (2008): "Eliciting Subjective Probabilities in Internet Surveys," *Public Opinion Quarterly*, 72(5), 866–891.

Dominitz, J., and C. F. Manski (1996): "Eliciting Student Expectation of the Returns to Schooling," *The Journal of Human Resources*, 31(1), 1–26.

Gan, L., M. Hurd, and D. McFadden (2005): "Individual Subjective Survival Curves," in *Analyses in Economics of Aging*, ed. by D. A. Wise, vol. I, pp. 377–411. University of Chicago Press, Chicago.

Hall, A., and T. R. Johnson (1980): "The Determinants of Planned Retirement Age," *Industrial and Labor Relations Review*, 33(2), 241–254.

Hamermesh, D. S. (1985): "Expectations, Life Expectancy, and Economic Behavior," *The Quarterly Journal of Economics*, 100(2), 389–408.

Hurd, M. D. (1989): "Mortality Risk and Bequests," *Econometrica*, 57(4), 779–813.

——— (2009): "Subjective Probabilities in Household Surveys," *Annual Review of Economics*, 1(1), 543–564.

Hurd, M. D., and K. McGarry (2002): "The Predictive Validity of Subjective Probabilities of Survival," *The Economic Journal*, 112, 966–985.

Keynes, J. (1921): *Treatise on Probability*. Macmillan & Co., Ltd.

Kleinjans, K., and A. van Soest (2010): "Nonresponse and Focal Point Answers to Subjective Probability Questions," *IZA Discussion Paper*, (5272).

Lusardi, A., and O. Mitchell (2007): "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education," *Business Economics*, 42(1), 35–44.

Manski, C. F. (1990): "The Use of Intentions Data to Predict Behavior: A Best-Case Analysis," *Journal of the American Statistical Association*, 85(412), 934–940.

——— (1993): "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?," in *Studies of Supply and Demand in Higher Education*, ed. by C. T. Clotfelter, and M. Rothschild, pp. 43–60. National Bureau of Economic Research, Inc.

——— (2002): "Identification of Decision Rules in Experiments on Simple Games of Proposal and Response," *European Economic Review*, 46, 880–891.

——— (2004): "Measuring Expectations," *Econometrica*, 72(5), 1329–1379.

Manski, C. F., and F. Molinari (2010): "Rounding Probabilistic Expectations in Surveys," *Journal of Business and Economic Statistics*, 28(2), 219–231.

Muth, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315–334.

Nyarko, Y., and A. Schotter (2002): "An Experimental Study of Belief Learning Using Elicited Beliefs," *Econometrica*, 70(3), 971–1005.

Pesaran, M. H. (1987): *The Limits to Rational Expectations*. Basil Blackwell, Oxford, U.K.

Ramsey, F. (1926): "Truth and Probability," in *1931, The Foundations of Mathematics and other Logical Essays*, ed. by R. Braithwaite, no. 1926, chap. VII, pp. 156–198. Kegan, Paul, Trench, Trubner & Co., London.

Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65(4), 781–831.

Part I

IN WHICH I DISCUSS SOME ECONOMETRIC IMPLICATIONS AND TREATMENTS OF SUBJECTIVE EXPECTATIONS

2 FLEXIBLE APPROXIMATION OF SUBJECTIVE EXPECTATIONS USING PROBABILITY QUESTIONS

This chapter is the reproduction of a paper written with Charles Bellemare and Sabine Kröger and forthcoming in the *Journal of Business and Economic Statistics*.

2.1 INTRODUCTION

The measurement of subjective expectations has proven useful for eliciting knowledge of economic agents and experts on the future realization of various economic variables (e.g., Dominitz and Manski (1997); Engelberg, Manski, and Williams (2009)) and improving the empirical content of stochastic models of choice under uncertainty (Delavande (2008); Bellemare, Kröger, and van Soest (2008)). It has been advocated that the measurement of expectations should proceed by first measuring subjective probability distributions. In particular, there is growing evidence that agents reveal different points of their subjective distribution (mean, median, or other quantiles) when asked for their best point prediction of a future event (see Manski (2004) for a review). Thus, deriving expectations from probability distributions can improve interpersonal comparisons while providing more information on the uncertainty faced by respondents.

Up to now, two approaches have been used to make inferences on subjective distributions. The first approach is parametric and assumes that the subjective distribution of a respondent is drawn from a parametric distribution (e.g., a normal or lognormal distribution) that depends on a finite number of unknown parameters. As with most parametric approaches, misspecification of the underlying distribution may lead to biased forecasts and inferences. The second approach is fully nonparametric, placing no restriction of the nature or shape of subjective distributions. This approach overcomes potentials biases due to misspecification of the underlying distribution at the expense of providing set rather than point identification of the functionals of interest.

In this paper, we present a flexible method that yields point identification of the distribution function of a respondent while maintaining weak assumptions on the shape of the underlying distribution. The flexible approach builds on cubic spline interpolation, which requires only that the underlying distributions be twice differentiable on their support. Moreover, the estimation by cubic splines involves solving a system of linear equations. Thus our flexible approach provides a simple analytical solution for the estimated function. Cubic splines are well-known interpolation methods (see e.g., Judd (1998)); however, to the best of our knowledge, they have not been applied to fit individual specific cumulative distribution functions using subjective expectations data. The closest work using interpolation methods to fit a cumulative distribution function is that of Kriström (1990), who estimated the population-level distribution of willingness to pay for an environmental good using linear interpolation and aggregated survey responses to valuation questions.

We illustrate our approach by revisiting the determinants of expectations concerning future income using data from the Survey of Economic Expectations (SEE). These data are characterized by high levels of censoring and potential rounding. Censoring occurs when individuals report a nonzero probability that the future outcome will fall outside the range of potential values spanned by the probability questions. The parametric approach maintains sufficiently strong distributional assumptions to deal with censoring. In contrast, the flexible approach maintains weaker distributional assumptions. As a result, estimated moments will be affected by censoring. To overcome this problem, we propose a quantile-based flexible approach that uses the estimated median as a measure of central tendency and the estimated interquartile range (IQR) as a measure of respondent uncertainty. We compare estimators of the determinants of expectations and uncertainty using both a specific parametric approach and our quantile-based flexible approach. We find that both approaches provide similar results for most determinants of future income, suggesting that the distributional assumptions chosen to implement the parametric approach are reasonable.

In the final part of the article, we present a Monte Carlo analysis designed to measure the impact of censoring and rounding on estimates of the determinants of expectations. We focus on comparing the performance of our flexible approach with that of a correctly specified parametric approach as well as an incorrectly specified parametric approach. We find that the flexible approach generates unbiased estimates of the determinants of expectations. This result holds when we introduce censoring and rounding levels believed to be present in the data. Moreover, the performance of the flexible approach is comparable to that of the correctly specified parametric approach but clearly outperforms the incorrectly specified parametric approach that we consider.

2.2 A Flexible Approach

Our objective is to approximate the subjective probability distribution $F_i(z) = \Pr_i(Z \leq z)$ of a respondent i using his or her answers to J probability questions of the type "what is the percent chance that Z

is less than or equal to z_j?", where $z_1 < z_2 < ... < z_J$ are threshold values. Thus the J data points available to make inferences on $F_i(z)$ are $\{(z_1, F_i(z_1)), ..., (z_J, F_i(z_J))\}$, where $0 \leq F_i(z_j) \leq 1$ denotes the probability statement to a question with threshold z_j. Censoring occurs when $F_i(z_1) > 0$ and/or $1 - F_i(z_J) > 0$. This implies that some probability mass is not contained within the interval $[z_1, z_J]$.

We propose to use the available data to approximate the subjective cumulative distribution function $F_i(z)$ using cubic spline interpolation. A cubic spline is a piecewise polynomial function defined on $J-1$ intervals, $[z_1, z_2]$,..., $[z_{J-1}, z_J]$. On each interval, the function $F_i(z)$ is assumed to be given by a polynomial $a_j + b_j z + c_j z^2 + d_j z^3$, where (a_j, b_j, c_j, d_j) are the interval-specific polynomial coefficients. The spline approximation of the function $F_i(z)$ is constructed by simply connecting the different polynomials at the relevant threshold values. The set $\{(a_j, b_j, c_j, d_j) : j = 1, \ldots, J-1\}$ contains the $4(J-1)$ unknown polynomial coefficients to be estimated. Exploiting continuity at the endpoints and interior thresholds provides $2J-2$ equations

$$\begin{aligned}
F_i(z_j) &= a_j + b_j z_j + c_j z_j^2 + d_j z_j^3 && \text{for } j = 2, \ldots, J-1 \\
F_i(z_j) &= a_{j+1} + b_{j+1} z_j + c_{j+1} z_j^2 + d_{j+1} z_j^3 && \text{for } j = 2, \ldots, J-1 \\
F_i(z_1) &= a_1 + b_1 z_1 + c_1 z_1^2 + d_1 z_1^3 \\
F_i(z_J) &= a_{J-1} + b_{J-1} z_J + c_{J-1} z_J^2 + d_{J-1} z_J^3
\end{aligned}$$

Next, restrictions that the first and second derivatives of $F_i(\cdot)$ agree at the interior thresholds generates $2J-4$ additional equations

$$\begin{aligned}
b_j + 2c_j z_j + 3d_j z_j^2 &= b_{j+1} + 2c_{j+1} z_j + 3d_{j+1} z_j^2 && \text{for } j = 2, \ldots, J-1 \\
2c_j + 6d_j z_j &= 2c_{j+1} + 6d_{j+1} z_j && \text{for } j = 2, \ldots, J-1.
\end{aligned}$$

Two more conditions, so-called "boundary conditions" at the endpoints, are needed to estimate the $4(J-1)$ polynomial coefficients of the cubic spline. There is very little guidance in the literature to chose these boundary conditions. Here we chose to impose that $F_i''(z_1) = F_i''(z_J) = 0$, yielding what is known in the literature as a natural cubic spline (see Judd (1998)). Thus restrictions on the derivatives and the boundary conditions generate a system of $4(J-1)$ linear equations that can be solved for the $4(J-1)$ unknown parameters. We experimented with boundary conditions restricting the first derivative at both endpoints $F_i'(z_1) = F_i'(z_J) = 0$ or by mixing restrictions on first and second derivatives (e.g., setting $F_i''(z_1) = F_i'(z_J) = 0$ or $F_i'(z_1) = F_i''(z_J) = 0$). We found that these changes had only minor effects on the estimated splines. We also experimented with linear and quadratic splines and found the cubic spline approximation to be superior. We did not find that increasing the order of the spline further increased the quality of the approximation. Thus, we use natural cubic splines throughout the rest of the article.

Absent censoring, moments can be directly estimated from the fitted subjective cumulative distribution function. In particular, the λ-th noncentral moment of Z can be computed analytically using

$$\widehat{\mathbf{E}}_i(Z^\lambda) = \sum_{j=1}^{J-1} \left[\frac{\widehat{b}_j z^{\lambda+1}}{\lambda+1} + \frac{2\widehat{c}_j z^{\lambda+2}}{\lambda+2} + \frac{3\widehat{d}_j z^{\lambda+3}}{\lambda+3} \Bigg|_{z_j}^{z_{j+1}} \right] \tag{2.1}$$

Approximating $\mathbf{E}_i(h(Z)) = \int h(z)dF_i(z)$ of a general function $h(\cdot)$ is slightly more complicated. In such cases, numerical integration can be performed by quadrature or simulation using $\widehat{F}_i(z)$. Similarly, quantiles can be obtained numerically by inverting $\widehat{F}_i(z)$. Quantiles are especially useful in the presence of censoring, which occurs when survey respondents report a nonzero probability that Z will fall below z_1 and/or above z_J. In such cases, relevant medians can be used as a measure of central tendency, and the interquartile range (IQR) can be computed as a measure of subjective uncertainty as long as $F_i(z_1)$ and $1 - F_i(z_J)$ are less than or equal to 0.25.

We illustrate the flexible approach by fitting three different distributions: a symmetric standard normal, an asymmetric chi-squared distribution with 3 degrees of freedom, and a bimodal distribution (with modes at $\pi/2$ and $5\pi/2$). The density of the bimodal distribution is given by $\frac{sin(z)+1}{A}$ over the [0,3π] interval, where $A = 2 + 3\pi$ ensures that the function integrates to 1 over its domain. We fitted each cumulative distribution function using between four and six data points equally spaced between 3 and -3 for the normal distribution, between 0 and 8 for the chi-squared distribution, and between 0 and 3π for the bimodal distribution. The results are reported in Figure 2.1. As expected, the goodness of fit increases with the number of data points for all three interpolations. A slight approximation error remains in the lower hand of the distribution when the number of data points is increased from four to six. Finally, we find that the approach has more difficulty fitting the bimodal distribution than the other two distributions. In contrast, the interpolation manages to provide a very good fit of the distribution with five or more data points.

Monotonicity

Cubic spline interpolation can produce oscillations that can cause the estimated distribution function to be nonmonotonically increasing. This is particularly problematic when estimating quantiles by inverting $\widehat{F}_i(\cdot)$ to obtain a unique solution. Perhaps the simplest and most effective way to correct for these oscillations is to use the Hyman filter (Hyman (1983)). This filter works in two steps. In a first step, define $\widehat{f}_i'(z_j)$ as the estimated value of the first derivative of the spline function at the threshold z_j. Next, define $S_{i-1/2} = (\widehat{F}_i(z_j) - \widehat{F}_i(z_{j-1}))/(z_j - z_{j-1})$ and $S_{i+1/2} = (\widehat{F}_i(z_{j+1}) - \widehat{F}_i(z_j))/(z_{j+1} - z_j)$ as the left-side slope con-

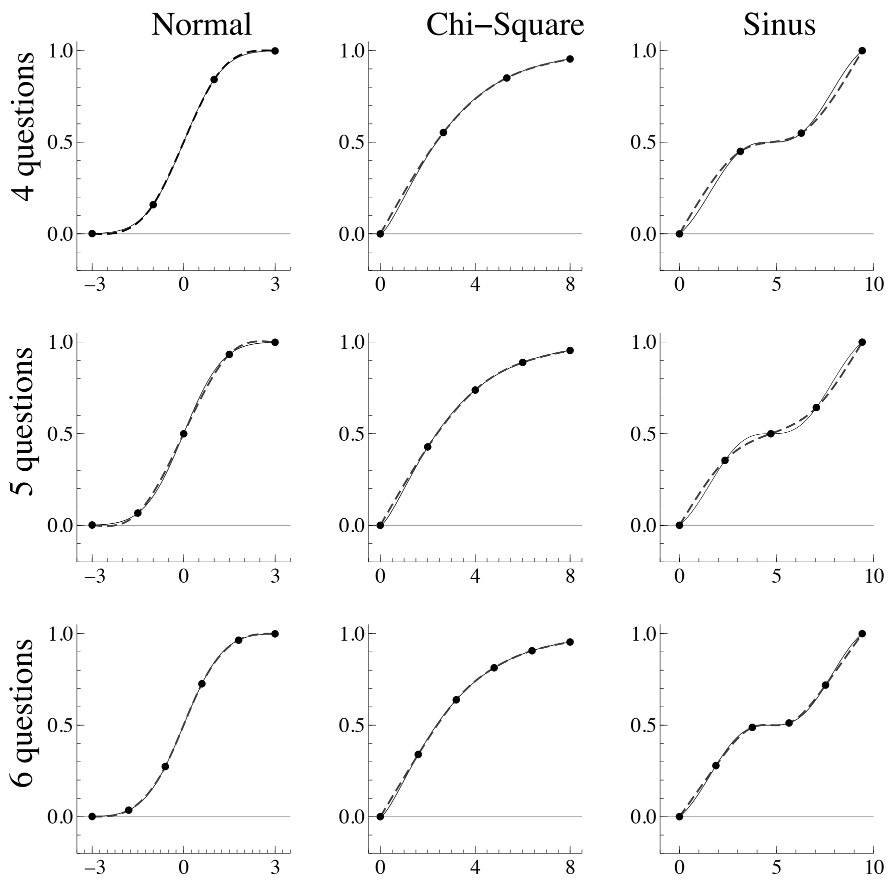


Figure 2.1: Fitted normal, chi-squared and sinus distributions using cubic spline interpolations with four to six data points (questions). The solid lines represent the true distributions. The dashed lines represent the fitted distributions using the data points (dark points).

necting with the previous threshold, $(\widehat{F}_i(z_{j-1}), z_{j-1})$, and the right-side slope connecting with the threshold $(\widehat{F}_i(z_{j+1}), z_{j+1})$. Boor and Swartz (1977) have shown that if an estimated function satisfies the criteria

$$0 \leq \widehat{f}_i'(z_j) \leq 3\min(S_{i-1/2}, S_{i+1/2}), \tag{2.2}$$

then it is monotone on the interval $[z_j, z_{j+1}]$. Thus the criteria (2.2) can be used to identify all points where the monotonicity condition is violated. In a second step, the condition of the equality of the second derivatives at each of the thresholds where monotonicity is violated is replaced by

$$\widehat{f}_i'(z_j) = \min\left[\max\left(0, \widehat{f}_i'(z_j)\right), 3\min\left(S_{i-1/2}, S_{i+1/2}\right)\right].$$

Hyman (1983) compared his filter approach to correct for nonmonotonicity with various alternative spline methods (e.g., Akima splines) and found that cubic spline interpolation coupled with his filter is the most effective method (in a mean squared error sense) to impose monotonicity on an estimated function.

2.3 Revisiting Expectations of Future Income

In this section we illustrate the flexible approach by revisiting data on income expectations that were previously analyzed in a parametric setting by Dominitz (2001). Data are taken from the 1994-1995 SEE administered through WISCON, a national telephone survey conducted by the University of Wisconsin Survey Center. We focus on the following survey question:

> *What do you think is the percent chance (or chances out of 100) that your own total income, before taxes, will be under \$$z_j$ (in the next 12 months)?*

For each respondent, four initial thresholds z_j were selected based on self-reported minimal and maximal values for their income support. Respondents could then be asked one or two additional questions based on their four answers. A detailed description of the branching algorithm to determine the income level or additional questions was presented by Dominitz (2001). We observe between four and six data points for each of 1,249 respondents in the SEE aged 25-59 who were active in the labor force at the time they answered the SEE and who provided all of the information for our analysis.

Figure 2.2 documents the extent of censoring in these data by plotting the sample distributions of $F_i(z_{1,i})$ and $F_i(z_{J,i})$. We find that only 44% of respondents have uncensored distributions at the lower end ($F_i(z_{1,i}) = 0$ in the left panel), whereas 66% of respondents have uncensored distributions at the upper hand ($F_i(z_{J,i}) = 1$ in the right panel). Only 37% of all sample respondents have uncensored distributions at both ends, a proportion too low to perform meaningful inferences using predicted moments. We deal with censoring by using the median as the measure of central tendency and the IQR as the measure of dispersion. Note that a small subsample of respondents have $F_i(z_{1,i})$ or $1 - F_i(z_{J,i})$ exceeding 0.25 and (to a lesser extent) exceeding 0.5; thus the estimated medians and/or IQR of respondents in this subsample are potentially biased. We report a Monte Carlo analysis to assess how such biases affect the measurement of the determinants of expectations.

We compared estimates using our proposed quantile-based flexible approach with those of a parametric approach applied to the same data. The parametric approach involves fitting the best lognormal distribution when sufficient data points are available. Respondents who state at most one value of $F_i(z_{j,i})$ that differs from 0 or 1 are fitted

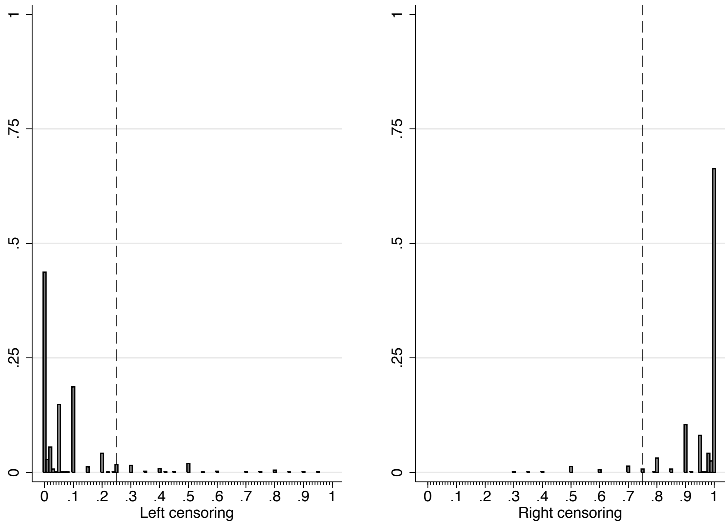


Figure 2.2: Distribution of $F_i(z_{1,i})$ (left) and $F_i(z_{J,i})$ (right) in the SEE data (N=1249). Dashed lines are at 0.25 (left) and 0.75 (right).

with the best *log-triangular* distribution, following the procedure of Engelberg, Manski, and Williams (2009).

We applied the Hyman filter for 850 respondents (68%) to correct for nonmonotonicity of the cumulative distribution function predicted by the flexible approach. Figure 2.3 presents a scatterplot of the predicted medians (left panel) and IQR (right panel) using both approaches (the flexible on the horizontal axis and the parametric on the vertical axis). We found similar predicted medians for both approaches, with predictions scattering closely and relatively equally below and above the 45-degree line. More important differences emerge when looking at the predicted IQR. There the flexible method tends to predict higher dispersions (74.5% of predicted IQR are below the 45-degree line).

We next estimated linear models for both approaches using the predicted medians and IQR as dependent variables and using a set of independent variables including realized income in the last year, basic demographic characteristics, employment status, and education level (using no high school diploma as the reference class). The first column of Table 2.1 presents some sample descriptive statistics of these variables. We estimated our models using the ordinary least squares (OLS) estimator with robust standard errors. Results are presented in subsequent columns of Table 2.1.

Overall, inferences using the flexible and parametric approaches are similar, suggesting that the assumption of expected income following a lognormal distribution is reasonable. Only small differences emerge. For instance, the flexible approach predicts that women and Hispanics expect significantly lower average median future income. Both meth-

	Desc. Stat.	Median		IQR 75-25	
	Mean (S.D.)	Param.	Flexible	Param.	Flexible
Income	36.592	0.841***	0.828***	0.225***	0.295***
	(32.999)	(0.063)	(0.052)	(0.039)	(0.040)
Self-employed	0.123	4.201**	3.206*	12.910***	9.579***
	(0.329)	(2.019)	(1.887)	(2.446)	(1.364)
Currently unemployed	0.059	0.272	-0.336	-4.031***	-1.291
	(0.236)	(1.897)	(1.431)	(1.271)	(1.140)
Previously unemployed	0.117	-3.956***	-3.532***	2.332	4.137***
	(0.321)	(1.315)	(1.193)	(1.694)	(1.446)
Female	0.467	-2.217	-2.756**	0.808	0.474
	(0.499)	(1.360)	(1.246)	(1.301)	(0.820)
Partner	0.651	0.440	1.226	-1.909	-0.357
	(0.477)	(0.867)	(0.757)	(1.421)	(0.720)
Age	39.001	-0.008	0.036	-0.216***	-0.159***
	(9.155)	(0.049)	(0.041)	(0.068)	(0.045)
White	0.877	-0.113	-0.340	-0.415	-1.624
	(0.329)	(1.539)	(1.459)	(1.659)	(1.360)
Black	0.067	0.516	-0.933	8.640**	2.222
	(0.251)	(2.258)	(2.001)	(4.227)	(1.685)
Hispanic	0.031	-3.080	-4.672**	4.296	-0.529
	(0.174)	(2.475)	(2.088)	(6.392)	(2.182)
High school diploma	0.164	-0.746	-0.813	-3.205	0.149
	(0.371)	(2.400)	(2.306)	(4.290)	(1.428)
Att. college w/o graduating	0.431	0.859	0.871	-2.028	0.619
	(0.495)	(2.344)	(2.255)	(4.202)	(1.355)
College graduate	0.365	5.139**	5.309**	-2.649	-1.326
	(0.482)	(2.598)	(2.449)	(4.172)	(1.534)
Constant		7.454**	6.117**	13.626***	7.019***
		(3.183)	(3.046)	(4.972)	(2.277)
R^2		0.775	0.804	0.146	0.421
N		1,249	1,249	1,249	1,249

Standard errors in parentheses (Eicker-White used in OLS estimation).

* Significant at 10% level

** Significant at 5% level

*** Significant at 1% level

Table 2.1: Determinants of subjective medians and IQR in the SEE using the parametric and flexible approaches

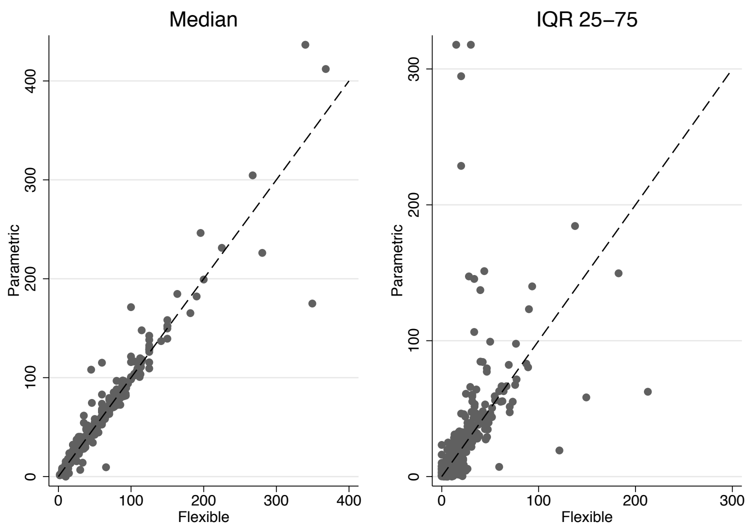


Figure 2.3: Scatterplot of estimated medians (left) and IQRs (right) of subjective income expectations, with either the parametric (vertical axis) or flexible (horizontal axis) method (N=1249). The dashed line represents the 45-degree line.

ods yield different results concerning the effect of unemployment on the uncertainty of future income and the income uncertainty faced by African-Americans. Results of the parametric approach suggest that the currently unemployed face significantly lower income uncertainty, whereas results of the flexible approach indicate that the previously unemployed have significantly higher income uncertainty. The parametric approach finds that African-American respondents have significantly greater income uncertainty, whereas this effect is smaller and insignificant using the flexible approach.

2.3.1 *Monte Carlo analysis*

We conducted a Monte Carlo analysis to assess how censoring and possible rounding in the SEE income data can affect the results in our application. Out analysis focuses on comparing the performance of our proposed flexible approach with the performance of the parametric approach, using both correctly specified and misspecified distribution functions for the parametric approach. We begin by specifying the data-generating process of medians med_i and interquartile ranges IQR_i

$$med_i = \theta_0 + \theta_1 x_{1i} + \varepsilon_i \tag{2.3}$$

$$IQR_i = \gamma_0 + \gamma_1 x_{2i} + \eta_i, \tag{2.4}$$

where x_{1i} and x_{2i} are two determinants, and where ε_i and η_i denote homoscedastic unobserved heterogeneity. Our objective is to analyze the properties of the OLS estimator of $(\theta_0, \theta_1, \gamma_0, \gamma_1)'$ in the presence of censoring and rounding. To proceed, we specify (2.3) and (2.4) as equations generating quantiles of a Kumaraswamy distribution defined over the [0,1] interval with parameters $(\alpha_i \geq 0, \beta_i \geq 0)$. The Kumaraswamy distribution is sufficiently flexible to accommodate a wide range of symmetric and asymmetric distributions of potential outcomes (Kumaraswamy (1980)). For example, $(\alpha_i = 2, \beta_i = 2)$ implies a symmetric distribution centered at 0.5, whereas $(\alpha_i = 1, \beta_i = 5)$ produces a severely left-skewed distribution with mode at 0.2. We specify our data-generating process in the following way. First, values of x_{1i} and x_{2i} are drawn from a uniform distribution on the [-0.5,0.5] interval, whereas values of ε_i and η_i are each drawn from a standard normal distribution with mean 0 truncated to the [-0.1,0.1] interval. Finally, we set $(\theta_0 = 0.5, \theta_1 = 0.3, \gamma_0 = 0.5, \gamma_1 = 0.3,)$. These data-generating processes force both med_i and IQR_i to lie within $[0.25, 0.75]$. We next present in detail the steps performed in our Monte Carlo simulations. Our analysis of the flexible and parametric approaches differs only with respect to step 4.

Step 1. Draw (med_i, IQR_i) for $i = 1, 2, ..., N$ using eqs. (2.3) and (2.4).

Step 2. Compute for each i the parameters (α_i, β_i) corresponding Kumaraswamy distribution by numerically solving the following system of equations:

$$\begin{aligned} med_i &= Q_{0.5}(\alpha_i, \beta_i) && (2.5) \\ IQR_i &= Q_{0.75}(\alpha_i, \beta_i) - Q_{0.25}(\alpha_i, \beta_i) && (2.6) \end{aligned}$$

such that $Q_\kappa(\alpha_i, \beta_i) = F^{-1}(\kappa; \alpha_i, \beta_i)$ where $F^{-1}(\cdot)$ denotes the inverse mapping of the Kumaraswamy cumulative distribution function $F(x) = 1 - (1 - x^{\alpha_i})^{\beta_i}$ evaluated at $0 \leq \kappa \leq 1$ with parameters (α_i, β_i).

Step 3. Generate points $\{z_{j,i} : j = 2, ..., J-1\}$ using a branching algorithm inspired by our empirical application. In particular, respondents with $med_i \leq 0.42$ are assigned the vector of thresholds (0,0.125,0.25,0.4,0.7,1), those with $0.42 < med_i < 0.59$ are assigned thresholds (0,0.2,0.4,0.6,0.8,1), and those with $med_i \geq 0.59$ are assigned thresholds (0,0.3,0.6,0.75,0.875,1). As with our empirical application, this algorithm assumes that prior information about the location of the distribution is used to generate thresholds. Then the cumulative probabilities $F(z_{j,i}; \alpha_i, \beta_i)$ are computed at all $z_{j,i}$ values.

Step 4 (flexible approach). Compute estimates $\widehat{med}_i$ and $\widehat{IQR}_i$ using the flexible approach.

Step 4 (parametric approach). Compute the value of δ that minimizes the following loss function:

$$\widehat{\delta} = \arg\min_{\delta} \sum \left(\Pr(Z \leq z_{j,i}; \delta) - F_i(z_{j,i}, r)\right)^2,$$

where the summation is over the data points of respondent i and $\Pr(Z \leq z_{j,i}; \delta)$ denotes a parametric cumulative distribution function with an unknown vector of parameters δ. We consider the correctly specified case where $\Pr(Z \leq z_{j,i}; \delta)$ is correctly chosen to be the Kumaraswamy distribution with parameters $\delta = [\alpha_i, \beta_i]$. We also consider a misspecified case where $\Pr(Y \leq z_{j,i}; \delta)$ is chosen to be the Normal distribution with mean τ_i and variance γ_i^2. We compute estimates $\widehat{med}_i$ and $\widehat{IQR}_i$ using $\widehat{\delta}$.

Step 5. Estimate the following equations:

$$\widehat{med}_i = \theta_0 + \theta_1 x_i + \bar{\varepsilon}_i \tag{2.7}$$

$$\widehat{IQR}_i = \gamma_0 + \gamma_1 x_i + \bar{\eta}_i, \tag{2.8}$$

where $\bar{\varepsilon}_i = \varepsilon_i + \widehat{med}_i - med_i$ and $\bar{\eta}_i = \eta_i + \widehat{IQR}_i - IQR_i$. Equations (2.7) and (2.8) are identical to eqs. (2.3) and (2.4), except that the true medians and IQRs are replaced by approximated values generated using either the parametric approach or the flexible approach. Estimated values $\left(\widehat{\theta}_0, \widehat{\theta}_1\right)'$ and $(\widehat{\gamma}_0, \widehat{\gamma}_1)'$ are saved. We repeat steps 1-5 for 10,000 samples of size 100.

The foregoing five steps generate our baseline results without censoring or rounding. To analyze the effects of rounding, we replace the probabilities $F\left(z_j; \alpha_i, \beta_i\right)$ in step 3 by the closest of the following numbers: 0, 1, 2, 3, 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 65, 70, 75, 80, 85, 90, 95, 97, 98, 99, or 100. This sequence closely matches the probability responses in our application. It also is one of the main rounding patterns discussed in the literature (see Manski and Molinari (2010)).

To analyze the effects of censoring, we randomly draw for each i a pair of censoring levels from below and from above using the empirical distribution of censoring levels presented in Figure 2.2. Let c_i^0 and c_i^1 denote these censoring levels. We then rescale the thresholds assigned in step 3 such that $z_{1,i} = Q_{c_i^0}(\alpha_i, \beta_i)$ and $z_{J,i} = Q_{1-c_i^1}(\alpha_i, \beta_i)$.

We evaluate the performance of the flexible and parametric approaches with rounding and censoring by computing parameter and standard error biases. Parameter bias is computed using $(\frac{1}{S}\sum_{s=1}^{S}\widehat{\phi}^s - \phi)/\phi$, where $\phi \in \{\theta_0, \theta_1, \gamma_0, \gamma_1\}$ are the true values and $\widehat{\phi}^s$ denotes the estimated parameter in simulation $s \leq S = 10,000$. We also compute the percent bias of the estimated standard errors using $(\frac{1}{S}\sum_{s=1}^{S} se(\widehat{\phi}^s) - sd(\widehat{\phi}^s))/sd(\widehat{\phi}^s)$, where $sd(\widehat{\phi}^s)$ denotes the standard deviation of all $\widehat{\phi}^s$ and $se(\widehat{\phi}^s)$ denotes the standard error predicted using the covariance matrix of the OLS estimator with homoscedasticity ($\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$). Thus we report the percent difference between the average standard error predicted by the OLS estimator and the actual standard deviation of the estimates over the 10,000 simulations.

Table 2.2 presents the results. We see that under the baseline scenario (no censoring or rounding), both parameter and standard error biases are small and negligible for the flexible and correctly specified parametric approaches. Of note, these results also hold when censoring and rounding levels believed to be present in our data are incorporated in the analysis. This suggests that results of our empirical application are robust to censoring and possible rounding in the data. We also find that our flexible approach clearly outperforms the misspecified parametric approach based on the erroneous assumption that distributions are normal in the population. There parameter bias is substantial: -24% for $\widehat{\theta}_1$, -17% for $\widehat{\gamma}_0$, and -43% for $\widehat{\gamma}_1$. These biases are not affected by censoring and rounding.

2.4 Conclusion and Discussion

Our Monte Carlo analysis suggests that the quantile-based flexible approach is robust to levels of rounding discussed in the literature and can accommodate censoring levels present in our data. We found that the flexible approach is comparable to a (first-best) correctly specified parametric approach in terms of bias and efficiency. Moreover, it clearly outperforms the misspecified parametric approach that we consider. We interpret these results as an indication that the flexible approach represents a potentially useful alternative to the existing parametric approach when researchers have little prior knowledge of the shape of the underlying distributions.

The flexible approach has three limitations. First, it lacks a distribution theory which would allow one to make inferences on individual specific distribution functions. This limitation might not pose a significant problem in practice, given that research on subjective expectations has focused on making statistical inferences on the determinants on expectations rather than on individual distribution functions. A second limitation is that moments are biased in the presence of censoring. This is expected because the flexible approach maintains weak assumptions on the shape of the distribution, thereby preventing extrapolation outside of the support spanned by the probability questions. Finally, our quantile-based flexible approach can accommodate only moderate levels of censoring.

Greater levels of censoring can be dealt with in several ways. The first and simplest way is to drop observations with excessive censoring. Though simple, this approach may introduce selection biases if the observations dropped represent a nonrandom subset of observations. A second way is to revert back to the parametric approach and maintain stronger distributional assumptions. Although this would allow accounting for censoring in the data, adopting a fully parametric approach introduces possible specification biases. Our analysis suggests that such biases can be sizeable. Finally, the survey design could be improved by designing probability questions to gather information on a larger range of possible outcomes. The flexible approach could then be used to make inferences while maintaining weaker assumptions on the underlying distributions.

	Flexible			Correct Parametric			Misspecified Parametric		
	Baseline	Rounding	Censoring	Baseline	Rounding	Censoring	Baseline	Rounding	Censoring
$\widehat{\theta}_0$	0.000	-0.000	-0.000	0.000	0.000	-0.000	-0.001	-0.001	-0.002
$\widehat{\theta}_1$	-0.008	-0.002	-0.007	0.001	0.001	-0.001	-0.241	-0.240	-0.223
$\widehat{\gamma}_0$	-0.010	-0.013	-0.008	0.000	-0.002	-0.000	-0.171	-0.173	-0.146
$\widehat{\gamma}_1$	-0.011	-0.016	-0.018	-0.000	0.001	-0.000	-0.430	-0.431	-0.367
std($\widehat{\theta}_0$)	-0.002	0.007	-0.011	-0.002	0.002	-0.011	-0.003	0.004	-0.013
std($\widehat{\theta}_1$)	-0.011	-0.005	-0.011	-0.015	-0.014	-0.014	-0.002	-0.003	-0.009
std($\widehat{\gamma}_0$)	0.003	-0.006	0.005	0.001	-0.007	0.004	0.002	-0.005	0.003
std($\widehat{\gamma}_1$)	0.040	0.046	0.035	-0.004	-0.002	-0.001	-0.027	-0.020	-0.030

Table 2.2: Lines $(\widehat{\theta}_0, \widehat{\theta}_1, \widehat{\gamma}_0, \widehat{\gamma}_1)$ present the corresponding parameter biases computed using $(\frac{1}{S}\sum_{s=1}^{S}\widehat{\phi}^s - \phi)/\phi$, where $\phi \in \{\gamma_0, \gamma_1, \theta_0, \theta_1\}$ are the true values and $\widehat{\phi}^s$ denotes the estimated parameter in simulation $s \leq S = 10,000$. Lines std$(\cdot)$ present the percent bias of the estimated standard errors of the corresponding estimated parameter using $(\frac{1}{S}\sum_{s=1}^{S} se(\widehat{\phi}^s) - sd(\widehat{\phi}^s))/sd(\widehat{\phi}^s)$, where $sd(\widehat{\phi}^s)$ denotes the standard deviation of all $\widehat{\phi}^s$ and $se(\widehat{\phi}^s)$ denotes the standard error predicted using the covariance matrix of the OLS estimator with homoscedasticity $(\sigma^2(\mathbf{X}'\mathbf{X})^{-1})$.

2.5 Acknowledgements

The authors thank Arthur van Soest, participants at the ESA meetings in Nottingham and Tucson, two referees, an associate editor, and the editor for comments. Financial support was provided by FQRSC.

BIBLIOGRAPHY

Bellemare, C., S. Kröger, and A. van Soest (2008): "Measuring Inequity Aversion in a Heterogeneous Population using Experimental Decisions and Subjective Probabilities," *Econometrica*, 76, 815–839.

de Boor, C., and B. Swartz (1977): "Piecewise Monotone Interpolation," *Journal of Approximation Theory*, 21(4), 411–416.

Delavande, A. (2008): "Pill, Patch or Shot ? Subjective Expectations and Birth Control Choice," *International Economic Review*, 49, 999–1042.

Dominitz, J. (2001): "Estimation of Income Expectations Models Using Expectations and Realization Data," *Journal of Econometrics*, 102, 165–195.

Dominitz, J., and C. F. Manski (1997): "Using Expectations Data to Study Subjective Income Expectations," *Journal of the American Statistical Association*, 92, 855–867.

Engelberg, J., C. Manski, and J. Williams (2009): "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters," *Journal of Business and Economic Statistics*, 27, 30–41.

Hyman, J. M. (1983): "Accurate Monotonicity Preserving Cubic Interpolation," *SIAM Journal on Scientific Computing*, 4(4), 645–654.

Judd, K. L. (1998): *Numerical Methods in Economics*. MIT Press, Cambridge, MA, U.S.A.

Kristrom, B. (1990): "A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies," *Land Economics*, 66, 135–139.

Kumaraswamy, P. (1980): "A Generalized Probability Density Function for Double-Bounded Random Processes," *Journal of Hydrology*, 46, 79–88.

Manski, C. F. (2004): "Measuring Expectations," *Econometrica*, 72(5), 1329–1376.

Manski, C. F., and F. Molinari (2010): "Rounding Probabilistic Responses in Surveys," *Journal of Business and Economic Statistics*, 28, 219–231.

3

BOUNDING PREFERENCE PARAMETERS UNDER DIFFERENT ASSUMPTIONS ABOUT BELIEFS: A PARTIAL IDENTIFICATION APPROACH

This chapter is the reproduction of a paper written with Charles Bellemare and Sabine Kröger and published in *Experimental Economics* (2010).

3.1 INTRODUCTION

A recent development in econometrics concerns the identification and estimation of econometric models that are partially identified (see Manski and Tamer (2002)). A model is partially identified if it maintains weaker assumptions than are necessary to point identify the parameters of interest. The approach allows researchers to understand what can be learned about a parameter of interest under different sets of assumptions, some potentially more plausible than others. Each set of assumptions can be used to place bounds around the model parameters of interest. These bounds in turn define the so-called identification region of the model parameters that contains all parameter vectors which are consistent with the data given the maintained assumptions. The identification regions can in turn be used to perform specification tests of the validity of maintaining stronger assumptions to point identify the model parameters. In particular, maintaining stronger but invalid assumptions concerning key variables may yield point estimates that fall outside the identification region derived under weaker assumptions.

Early applications have focused on placing bounds around moments or quantiles of a conditional distribution (see Manski (1989, 1994)) . These applications are non-parametric in nature: identification regions around moments or quantiles are estimated using the data alone without referring to a specific parametric model. More recently, the approach has been extended to make inferences on parameters of incomplete parametric and semi-parametric models (see Manski and Tamer (2002)). Applications of the later include Honoré and Tamer (2006) and Ciliberto and Tamer (2009). To our knowledge, these methods have yet to be applied to experimental data.

In this paper we illustrate the usefulness of these methods by making inferences on preferences in a choice problem with uncertainty under different assumptions about the beliefs of players.[1] More specifically,

1 This paper relates to two approaches used so far to separately identify the effects of preferences and beliefs on decision making under uncertainty. The first approach

we specify a simple model of sender behavior in a binary investment game (see Berg, Dickhaut, and McCabe (1995)). We model decisions of senders as a function of their expected final payoffs (which proxies their trust in the responder), a component capturing other-regarding preferences, and an unobserved random component. We focus on relating the size of the identification regions to the restrictiveness of the assumptions maintained on the beliefs of senders. We explore three different sets of assumptions. The first and weakest set of assumptions states that researchers have no information about beliefs of senders apart from the natural restrictions imposed by the game (e.g., the amount returned must be below and above known boundaries). The second set of assumptions states that all senders expect to receive not less when they invest than when they do not. This second set is more restrictive than the first. As a result, we expect the identification region under the second set to be contained in the identification region derived under the first set of assumptions. The third and most restrictive set of assumptions we consider consists of assuming that senders have rational expectations. We show that the latter set of assumptions produce the smallest identification region of the three we consider. Finally, we point estimate our model parameters using non-incentivized beliefs stated by senders in the experiment. Our point estimates suggests that expectations about responder behavior as well as other-regarding preferences are both significant determinants of investments. Moreover, we find that our point estimates fall within the first two identification regions. This suggests that reasonable inferences on preferences can be obtained using non-incentivized beliefs.

The rest of the paper is organized as follows. Section 3.2 presents the experimental design and the data. Section 3.3 the econometric model. Section 3.4 presents our results. Section 3.5 concludes.

3.2 EXPERIMENTAL DESIGN AND PROCEDURE

3.2.1 *Experimental design*

Our experimental design is a modified version of the two player investment game of Berg, Dickhaut and McCabe (1995). In our experiment, senders and responders were both endowed with 6$US.[2] Contrary to Berg, Dickhaut and McCabe (1995), we restricted the decision space of senders to two choices: investing all or none of the endowment. If a sender invested his endowment, that amount was doubled and added to the endowment of the responder. In turn, the responder had the

compares behavior in treatments with uncertainty with behavior in treatments where uncertainty is blocked by design (see, e.g., Cox (2004)). The second approach uses data on subjective beliefs to recover estimates of preference parameters (see, e.g., Bellemare, Kröger, and van Soest (2008)).

2 The complete content of the computer screens can be downloaded from http://www.ecn.ulaval.ca/charles.bellemare/.

opportunity to return any amount from his augmented endowment to the sender (i.e., he could return up to 18$).[3] If the sender did not invest his endowment, the responder could return any amount from his initial endowment (up to 6$).

Responders made their decisions using the strategy method: they each had to decide how much to return when the sender invested his endowment, and how much to return when the sender would not invest his endowment. The decision that corresponded to the actual choice of the sender was chosen to be the effective action and determined the payoff of both participants. After making their decisions, senders were asked to state their subjective beliefs. Before stating their beliefs, they were further reminded of the decision tasks and given examples to clarify the belief elicitation procedure. Senders were not rewarded for the accuracy of their beliefs.

Senders had to state their subjective beliefs in two scenarios. They were first asked to state their beliefs if they did not invest. In particular, they had to state how many out of 100 responders would return 0$, and how many would return amounts in the following intervals $\{(0,1], (1,2], (2,3], (3,4], (4,5], (5,6]\}$.[4] By allowing senders to place a positive probability on getting back 0, we allow their subjective distribution functions to be censored from below. Additionally, senders were asked to state their beliefs about responder behavior if they invested their endowment. Senders were asked to state how many out of 100 responders would return 0$, and how many would return amounts in the following intervals $\{(0,3], (3,6], (6,9], (9,12], (12,15], (15,18]\}$.[5,6]

3.2.2 *Experimental procedure*

After all participants had made their decisions, senders and responders were randomly matched and payoffs were computed based on the decisions of the pair. Participants were then informed of the outcome of

3 Expending the choice set of senders is in principle possible, but this will require asking each participant to answer many more questions on their beliefs (see below).

4 If the probability mass entered exceeded 100, senders where automatically instructed to go back and adjust their answers.

5 In order to detect whether senders stated beliefs to rationalize their decisions, we randomized approximately one third of all participants in our experiment to a group of "observers" who did not make any decisions but who answered the belief questions after having read the same instructions as all other participants. Observers received each 6$ for their participation. We found no significant differences between the beliefs of senders and those of observers. See the extended working paper version of the paper for details (Bellemare, Bissonnette, and Kröger, 2007).

6 At the end of the experiment we elicited participants' risk preferences. We asked participants to play a sequence of lotteries similar to that proposed by Holt and Laury (2002). We will not discuss those results further as we found no significant relationship between measured risk preferences and investment behavior. Similar results have been reported by Eckel and Wilson (2004) and Houser, Schunk, and Winter (2010).

the experiment and their final payoffs. The experiment was conducted in May 2005 at the Economic Science Laboratory at the University of Arizona using the software zTree (Fischbacher (2007)). Participants were recruited via email and were mainly students in finance, business administration, economics, and engineering. Participants received a 5$ show-up fee upon arrival at the laboratory. We observed 38 pairs of players in 9 sessions of the experiment. An experimental session lasted on average 60 minutes, and, including their show up fee, participants earned on average 12.18$ (9.92$ for senders and 15.87$ for responders).

3.2.3 *Descriptive statistics*

24 of the 38 senders (63%) invested their endowment. To gain some insights on whether investors and non-investors trusted responders differently, we compare the subjective belief distributions of investors with those of non-investors. Figure 3.1 presents the average subjective belief distributions of investors (light bars, $N = 24$) and non investors (dark bars, $N = 14$). We find that both groups had similar beliefs about responder behavior if they consider not investing their endowment. In particular, both investors and non investors place on average a very high probability of getting nothing back from responders. In fact, we fail to find significant differences between the distribution of beliefs of investors and non-investors in each of the seven brackets of amounts reported in Figure 3.1.[7]

Differences between both investors and non-investors emerge when we look at their beliefs when investing their endowment. There, non-investors placed a 48.3% probability on getting nothing back from responders, substantially less than the 24.6% probability placed by investors. A Mann-Whitney U test easily rejects the null hypothesis that the distributions of beliefs about getting nothing back when investing are the same (p-value = 0.012). Moreover, Mann-Whitney U tests reject the null hypothesis that distributions of beliefs of investors and non-investors for the interval $(9, 12]$ are the same (p-value = 0.050). Together these results suggest that investors expect to get more when investing their endowment than non-investors.[8]

To assess whether the beliefs of senders were rational, we computed for each sender the deviation of their subjective expectations how much the responder would return when they would invest (when they would not invest) and the observed average amount returned for this case 0.26$ (observed average when not investing: 3.66$). Figure 3.2 presents the distributions of these differences. We find small discrepancies between expectations and observed responses when not

7 We tested for each interval $(0, (0, 1], \ldots)$ the null hypothesis the distributions of beliefs are the same for investors and non-investors using a Mann-Whitney U test. The lowest p-value out of the seven intervals tested is 0.238.

8 We do not find significant differences between the distributions of beliefs of both groups in the intervals $(0, 3]$, $(3, 6]$, $(6, 9]$, $(12, 15]$, and $(15, 18]$.

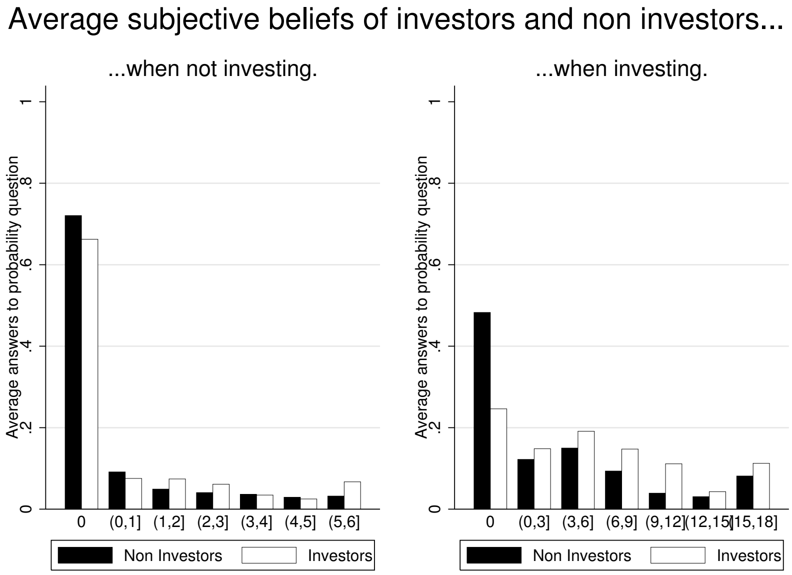


Figure 3.1: Subjective beliefs about the amount returned separately for investors (light bars, $N = 24$) and non investors (dark bars, $N = 14$) when not investing (left panel) and when investing (right panel).

investing, reflecting the fact that most senders correctly anticipated that the probability of getting close to nothing would be high when not investing. More substantial discrepancies emerge when considering amounts returned when investing. There, we find that a substantial amount of senders have expectations below and above the observed amount returned. Even though we fail to reject the Null hypothesis that the median deviation is equal to zero in both cases (p-value = 0.545 when not investing and 0.354 when investing), we find that the 25^{th} and 75^{th} percentiles of the distributions are significantly different from 0.[9] Deviations observed in Figure 3.2 may also reflect noise rather than genuine deviations from rational expectations. Separating noise from true underlying beliefs is out of the scope of the paper. However, if beliefs are mostly noise, they should be poorly related to decisions of senders. This issue is discussed in the next section.

3.3 A SIMPLE MODEL OF CHOICE

We assume that the utility of not investing for sender i is given by $u_i^{keep} = \beta(w + r_i^{keep})$, where r_i^{keep} denotes the amount the responder returns to sender i when i does not invest, w denotes the initial endowment of sender i, and β measures the marginal utility of income.

9 We reject the Null hypothesis that the deviation is equal to zero at the 25th and 75th percentiles for both scenarios (p-value=0.000 and 0.042 when not investing and p-value=0.020 and 0.029 when investing).

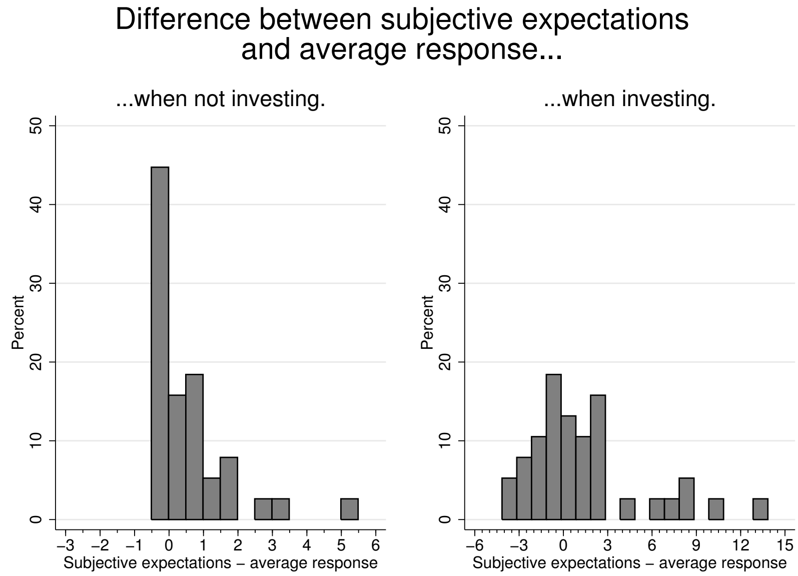


Figure 3.2: Distribution of the difference between subjective expectations of all senders and observed average response of all responders in the event of not investing (left graph, $N = 38$) and in the event of investing (right graph, $N = 38$).

The amount returned when not investing r_i^{keep} can vary between 0 and the endowment $w = 6\$$ of the responder.

When sender i invests, he foregoes his endowment w which is then doubled and transferred to the responder. As a result, a surplus of w is created when investing. We model the utility of investing as $u_i^{invest} = \beta r_i^{invest} + \theta$, where r_i^{invest} denotes the amount returned by the responder when investing,[10] and θ captures any utility gain coming from some form of other regarding preferences, whether it is a concern for efficiency or altruism.[11] Recent studies suggest that concerns for social efficiency may be particularly important (see Engelmann and Strobel (2004)). In terms of our model, this would imply that $\theta > 0$.

We next assume that senders make their decisions by comparing their subjective expected utilities of investing and not investing. The expected utilities of not investing and investing are given by

$$\mathbf{E}\left(u_i^{keep}\right) = \beta\left(w + \mathbf{E}\left(r_i^{keep}\right)\right) + \epsilon_i^{keep} \tag{3.1}$$

$$\mathbf{E}\left(u_i^{invest}\right) = \beta\mathbf{E}\left(r_i^{invest}\right) + \theta + \epsilon_i^{invest}, \tag{3.2}$$

10 The amount returned r_i^{invest} by the responder can take a value between 0 and $3w = 18\$$.

11 The preferences presented here are equivalent to linear altruism (e.g., Andreoni and Miller, 2002): $u_i = \alpha x_i + \gamma x_j$ with $\beta = \alpha - \gamma$, $\theta = \gamma \cdot 3w$ where $x_i = r^{invest}$ and $x_j = (3w - r^{invest})$ denote income of player i and j. For the case of not investing, $\gamma = 0$. Our data does not allow us to identify more general preferences (for instance as in Charness and Rabin (2002)).

where the expectations are computed with respect to the subjective distribution functions of sender i. To allow for the fact that some senders will make sub-optimal choices, we add standard normal error terms ϵ_i^{invest} and ϵ_i^{keep} to the true expected utilities $\mathbf{E}(u_i^{invest})$ and $\mathbf{E}(u_i^{keep})$, and assume that sender i chooses the option $j \in \{keep, invest\}$ that maximizes $\mathbf{E}(u_i^j) + \epsilon_i^j$ rather than $\mathbf{E}(u_i^j)$.

3.4 IDENTIFICATION REGIONS OF THE MODEL PARAMETERS

We first characterize the identification region of (β, θ) that is consistent with the observed choice distribution of senders without imposing any information on beliefs. To estimate this region, we first consider the extreme case where all senders expect to receive with probability 1 the highest possible amount when investing ($r^{invest} = 3w$) and the lowest possible amount when not investing ($r^{keep} = 0$). This gives rise to the largest payoff difference between investing and not investing. In this case, the decision rule is to invest when $\mathbf{E}\left(u_i^{invest}\right) > \mathbf{E}(u_i^{keep})$, or equivalently

$$\beta(2w) + \theta + \epsilon_i > 0. \tag{3.3}$$

where $\epsilon_i = \epsilon_i^{invest} - \epsilon_i^{keep}$. A second extreme case occurs when all senders expect to receive with probability 1 the lowest amount possible when investing ($r^{invest} = 0$), and the highest possible amount when they do not invest ($r^{keep} = w$). This gives rise to the smallest payoff difference between investing and not investing. In this case, senders i will invest when

$$\beta(-2w) + \theta + \epsilon_i > 0. \tag{3.4}$$

Assuming that errors ϵ_i are statistically independent of each other and follow a standard normal distribution, aggregating inequalities (3.3) and (3.4) across the population yields the following set of inequalities relating the population probability of investing to the model parameters

$$\Phi\left(\beta(-2w) + \theta\right) \leq \Pr(invest) \leq \Phi\left(\beta(2w) + \theta\right) \tag{3.5}$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution. The identification region for (β, θ) contains all vectors of parameters that satisfy inequalities (3.5).

The shaded area in Figure 3.3 represents the identification region estimated by replacing $\Pr(invest)$ with the proportion of investments observed in our sample. It is immediate from (3.5) that θ is point-identified and equal to $\Phi^{-1}(\Pr(invest))$ when expectations have no influence on the decision process ($\beta = 0$). Otherwise, the observed proportion of investments is compatible with any combination of $\beta > 0$ and θ within the shaded area. We can easily see that the identification region of the social preference parameter θ increases with β, the strength of the effect of expectations on investment behavior. A smaller identification region can be derived by assuming that all

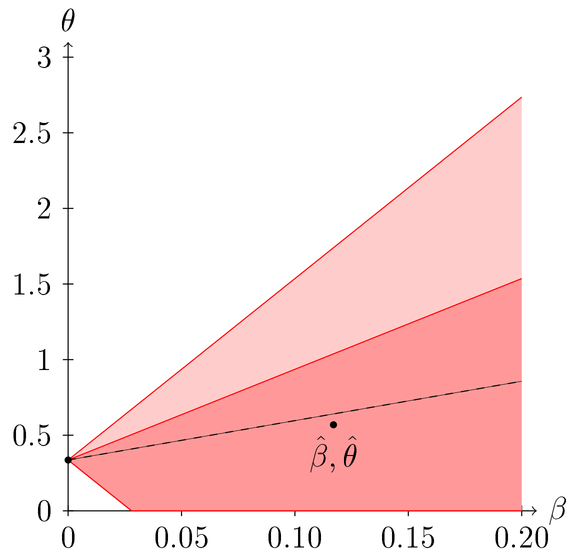


Figure 3.3: Estimated identification regions without information on subjective expectations (both shaded areas), assuming that $\mathbf{E}(r_i^{invest}) \geq \mathbf{E}(r_i^{keep})$ (dark shaded area only), and under rational expectations (dashed line). The point $(\hat{\beta}, \hat{\theta})$ denotes the parameter estimates obtained using the subjective expectations data.

senders expect to receive when they invest at least or more than when they do not ($\mathbf{E}(r_i^{invest}) \geq \mathbf{E}(r_i^{keep})$). Under this assumption, inequality (3.3) remains unchanged as it does not violate the new restriction on beliefs. Inequality (3.4) on the other hand concerns the lowest possible payoff difference, a difference of 0 under the new restriction ($\mathbf{E}(r_i^{invest}) = \mathbf{E}(r_i^{keep})$. In this case, senders i will invest when

$$\beta(-w) + \theta + \epsilon_i > 0. \tag{3.6}$$

Aggregating inequalities (3.3) and (3.6) across the population produces a new set of inequalities relating the population probability of investing and the model parameters

$$\Phi\left(\beta(-w) + \theta\right) \leq \Pr(invest) \leq \Phi\left(\beta(2w) + \theta\right). \tag{3.7}$$

The smaller identification region derived from (3.7) is given by the dark shaded area in the Figure 3.3. As expected, the new area is a strict subset of the area derived previously as it places much tighter upper bounds of the social preference parameter θ.

Another way to reduce the size of the identification region is to assume that senders have objectively correct (rational) expectations. This would imply that $\mathbf{E}\left(r_i^{invest}\right)$ and $\mathbf{E}\left(r_i^{keep}\right)$ both coincide with observed average responder behavior, $\overline{r}^{invest}$ and $\overline{r}^{keep}$, and are common for all players. Then, the identification region is a line, connecting all values of β and θ that solve

$$\Phi\left(\beta(\overline{r}^{invest} - \overline{r}^{keep}) - \beta w + \theta\right) = \Pr(invest). \tag{3.8}$$

The dashed straight line in Figure 3.3 represents the estimated identification region obtained under the assumption that beliefs are rational, estimated by replacing $\overline{r}^{invest}$ and $\overline{r}^{keep}$ with the corresponding sample averages. We see that the assumption of rational expectations does not point identify the model parameters. This follows because all players are assumed to have the same information set. Hence, there is no variation in beliefs across players that would be needed for the point-identification the model parameters.

In our experiment, however, participants have heterogeneous beliefs (see section 3.2.3). This fact not only contradicts the rational expectation hypothesis but can be exploited to point identify the parameters. To illustrate this, we finally estimate the parameters of our model using the beliefs stated by each sender. To proceed, we replaced the unknown expectations $\mathbf{E}(r_i^{keep})$ and $\mathbf{E}(r_i^{invest})$ in (3.1) and (3.2) with expectations approximated using the cubic spline interpolation method proposed in Bellemare, Bissonnette, and Kröger (forthcoming).[12] We find that the estimated value of β is 0.117 (standard error = 0.065) and is significant at the 5% level against the one-sided alternative that $\beta > 0$. This suggests that the marginal utility of income is greater than zero. This significant relation also suggests that non-incentivized subjective beliefs can be used to successfully predict behavior. We further find that the other-regarding preference parameter θ is 0.569 (standard error = 0.241) and significant at the 5% level against a two-sided alternative.[13] This suggests that social preferences play a significant role in determining investments in the game. Figure 3.3 plots this point estimate.

We find that the point estimate lies within the first two identifications regions. The first region was obtained by taking into account all the possible beliefs that respondents could have. Therefore, the point estimate will fall by construction within this zone. The point estimate could fall outside the second identification region if the beliefs of players systematically violated the maintained assumption on beliefs, i.e., that senders will not be worse off when investing ($\mathbf{E}(r_i^{invest}) \geq \mathbf{E}(r_i^{keep})$, used to derive the second identification region. In our data, however, all senders expect to receive from the responder at least as much if they send their endowment than if they keep it.

12 Cubic spline interpolation allows to approximate expectations with minimal assumptions concerning the shape of the underlying distributions. Bellemare, Bissonnette, and Kröger (forthcoming) show that the bias when approximating a subjective mean is negligible given the number of probability questions answered by each sender.

13 The standard errors are possibly a little conservative as they do not account for noise in the approximated expectations.

Finally, we see that the point estimate using subjective expectations data lies in close proximity to the dashed line representing the identified parameter combinations assuming rational expectations. Moreover, we do not find significant differences between the point estimate and the dashed line.[14] Section 3.2.3 revealed that the distribution of subjective beliefs is centered around the observed response behavior. This together with the fact that in the simple linear model used here to illustrate the partial identification approach senders' decisions are based on the mean of their subjective expectations probably explains why the dashed line and the point-estimate are close. A model that relies on the whole belief distribution, as for instance a model including risk aversion, would very likely lead to a greater difference between the inferences that one can draw using subjective vs. rational expectations.

3.5 CONCLUSION

In this paper we have discussed recent developments in the area of partial identification of econometric models using the stylized example of a binary investment game. We have shown how bounds around model parameters can be derived under various levels of assumptions concerning the beliefs of players. We have also shown how these bounds can be used to assess the validity of using data on beliefs collected without providing players incentives to report them truthfully. Our results provide support for eliciting non-incentivized subjective expectations data: point estimates using these belies fall within our most reasonable bounds. More importantly, this paper has highlighted how the partial identification approach can be used to make inferences in a parametric model under weak assumptions about the beliefs of players in the investment game.

Another particularly promising area of future research would be to ask what can be learned about the prevalence of belief dependent preferences such as reciprocity and guilt aversion without information on beliefs. Belief-dependent preferences typically involve second-order beliefs, that is beliefs of players over the distribution of beliefs of other players. Elicitation of second-order beliefs is complicated by several factors. First, the task is cognitively more demanding than collecting data on first-order beliefs. Second, consensus effects may lead to a spurious correlation between decisions and stated second-order beliefs,

14 We estimated by bootstrap the 95% confidence region around our point estimate as well as a 95% confidence region around the dashed line by bootstrap. In particular, we generated 1000 bootstrap samples, sampling with replacement the decision and beliefs of senders. We computed for each bootstrap sample the point estimate as well as the dashed line. Computing both estimates using the same samples allows us to control for the correlation between the estimated dashed line and the point estimates that both rely on the same data. We find that both confidence regions overlap substantially.

thus biasing the quantitative importance of these preferences (see eg., Ellingsen, Johannesson, Torsvik, and Tjøtta (2010), Bellemare, Sebald, and Strobel (2010)). The tools of partial identification may provide a way to learn about the relevance of these preferences while avoiding the potential problems posed by elicitation of second-order beliefs.

The application of partial identification analysis in experimental economics goes beyond the partial observability of player beliefs. For instance, in many common experiments, interval responses are elicited (as opposed to point-valuations) using multiple price lists, as discussed by Andersen, Harrison, Lau, and Rutström (2006). Multiple price lists are frequently used in experiments to measure preference parameters, willingness to pay, or discount rates. Interval regressions used to analyze interval responses elicited using multiple price lists typically impose sufficiently strong parametric assumptions on the distribution of unobservables to point estimate the model parameters (see eg. Coller and Williams (1999)). The tools of partial identification, on the other hand, allow researchers to bound the model parameters under minimal assumptions about the location of the true valuations within the intervals of each respondent. Manski and Tamer (2002) show how bounds around model parameters can be derived in this setting. The estimated bounds can thus be contrasted with point estimates obtained using stronger assumptions, thus providing a basis for model specification testing.

Finally, partial identification can also be useful to understand the preferences of players in games with multiple equilibria. Multiplicity of equilibria severely complicates point estimation of the heterogeneity in preferences of players. One way to point identify preferences has been to assume an equilibrium selection procedure (eg. randomly selecting one of the possible equilibriums). Ciliberto and Tamer (2009) show how bounds can be placed around the choice probabilities in discrete games without imposing any equilibrium selection procedure. As we have stressed in this paper, these bounds can then be used to perform meaningful inferences on the model parameters characterizing the decision rules of players in the game.

BIBLIOGRAPHY

Andersen, S., G. W. Harrison, M. I. Lau, and E. E. Rutström (2006): "Elicitation Using Multiple Price List Formats," *Experimental Economics*, 9, 383–405.

Andreoni, J., and J. Miller (2002): "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 70(2), 737–753.

Bellemare, C., L. Bissonnette, and S. Kröger (2007): "Flexible Approximation of Subjective Expectations Using Probability Questions: An Application to the Investment Game," *IZA Discussion paper 3121*.

——— (forthcoming): "Flexible Approximation of Subjective Expectations Using Probability Questions," *Journal of Business and Economic Statistics*.

Bellemare, C., S. Kröger, and A. van Soest (2008): "Measuring Inequity Aversion in a Heterogeneous Population using Experimental Decisions and Subjective Probabilities," *Econometrica*, 76, 815–839.

Bellemare, C., A. Sebald, and M. Strobel (2011): "Measuring the Willingness to Pay to Avoid Guilt: Estimation using Equilibrium and Stated Belief Models," *Journal of Applied Econometrics*, 26(3), 437–453.

Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10, 122–142.

Charness, G., and M. Rabin (2002): "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics*, 117, 817–869.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 1791–1828.

Coller, M., and M. Williams (1999): "Elicitating Individual Discount Rates," *Experimental Economics*, 2, 107–127.

Cox, J. C. (2004): "How to Identify Trust and Reciprocity," *Games and Economic Behavior*, 46, 260–281.

Eckel, C., and R. K. Wilson (2004): "Is Trust a Risky Decision ?," *Journal of Economic Behavior and Organization*, 55, 447–465.

Ellingsen, T., M. Johannesson, G. Torsvik, and S. Tjøtta (2010): "Testing Guilt Aversion," *Games and Economic Behavior*, 68, 95–107.

Engelmann, D., and M. Strobel (2004): "Inequality Aversion, Efficiency and Maximin Preferences in Simple Distribution Experiments," *American Economic Review*, 94(4), 857–869.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

Holt, C. A., and S. K. Laury (2002): "Risk Aversion and Incentive Effects," *American Economic Review*, 92(5), 1644–1655.

Honoré, B., and E. Tamer (2006): "Bounds on Parameters in Dynamic Discrete Choice Models," *Econometrica*, 74, 611–629.

Houser, D., D. Schunk, and J. Winter (2010): "Distinguishing Trust from Risk: An Anatomy of the Investment Game," *Journal of Economic Behavior and Organization*, 74(1-2), 72–81.

Manski, C. (1989): "Anatomy of the Selection Problem," *Journal of Human Resources*, 24(3), 343–360.

——— (1994): "The Selection Problem," in *Advances in Econometrics: Sixth World Congress*, ed. by C. Sims, vol. 1, pp. 143–170. Cambridge University Press, Cambridge, U.K.

Manski, C., and E. Tamer (2002): "Inference on Regressions with Interval Data on a Regressor or Outcome," *Econometrica*, 70, 519–546.

Part II

IN WHICH I DISCUSS PREFERENCES AND EXPECTATIONS CONCERNING PENSIONS AND SURVIVAL (OR THE LACK THEREOF)

4

THE FUTURE OF RETIREMENT AND THE PENSION SYSTEM: HOW THE PUBLIC'S EXPECTATIONS VARY OVER TIME AND ACROSS SOCIO-ECONOMIC GROUPS

The content of this chapter is based on joint work with Arthur van Soest that previously appeared in the *CentER Discussion Paper Series* (2011-065).

4.1 INTRODUCTION

In models of life cycle behavior and inter-temporal decision making under uncertainty, expectations play an important role. For example, consumption, saving, and labour supply decisions of individuals and households not only depend on their current tastes and opportunities, but also on their expectations of future prices, their future income, etc. (see, for example, Feldstein, 1974).

Future expectations often remain unobserved, and traditional macro- or micro-economic models typically make assumptions on how they are formed, e.g. assuming rational expectations. The conclusions from these models may be biased if the assumptions on expectations are not satisfied. To solve the problem that expectations are unobserved, many recent empirical studies aim at measuring expectations directly using survey questions. See, for example, Manski (2004) for an assessment of the validity of this approach. Examples are Keane and Runkle (1990) on inflation expectations, Dominitz and Manski (2005) on expectations of equity returns, Dominitz and Manski (1997), Das et al. (1999) and Dominitz (2001) on income expectations, Hurd and McGarry (1995) on length of life expectations, Stephens (2004) on job loss expectations, Benítez-Silva and Dwyer (2005) on retirement expectations, Dominitz and Manski (2006) and Chan and Stevens (2008) on pension expectations, and Delavande and Rohwedder (2008) on expectations of old age social security income.

Pension expectations have become particularly relevant since the aging of the population has led to a debate in many industrialized countries on the need for pension reforms to keep the pension system sustainable (see, for example, Lindbeck and Persson, 2003, and Zaidi, 2010). Particularly since the economic and financial crisis, high retirement replacement rates can no longer be taken for granted. Governments are trying to increase awareness of pension risks and individual responsibility to guarantee financial security after retirement.

In this paper, we analyze expectations of the Dutch population aged 25 and older concerning the future generosity of the two main pillars of the Dutch system of income provision to the elderly – old age social security benefits (AOW) and occupational pensions (mandatory for almost all employees). As in many other European countries, sustainability of income provision is old age has become an important issue in public policy discussions due to the aging of the population; see, e.g., Bovenberg and Gradus (2008). Generous early retirement benefits are gradually being phased out and replaced by actuarially fair flexible retirement systems and the idea of working after the normal retirement age of 65 years has slowly become a real option, although impediments remain (e.g., Van Solinge and Henkens, 2007). The debate has been reinforced by the financial and economic crisis, leading to additional pressure on old age social security due to government budget concerns and to pressure on occupational pensions caused by the reduced value of the assets of occupational pension funds, who have invested part of the pension savings of their clients in equity. This makes it particularly interesting to analyze how different socio-economic groups forecast the future of the Dutch pension system and whether and how these forecasts have changed during the recent years under the influence of the public discussion and the financial and economic crisis.

The first reason why we think studying subjective beliefs is important relates to economic modelling. Many studies in the past, some already reviewed above, disproved the rational expectation hypothesis as the heterogeneity in beliefs observed in elicited subjective expectations is at odds with the rationality hypothesis. In our sample, at each given point in time, all the respondents are asked to predict the same outcome and in principle have access to the same information, so that under the assumption of rational expectations, they should all come to the same conclusion. The large heterogeneity in reported expectations in our data reveals that they do not. One of the explanations for this heterogeneity could be that some groups lack the proper cognitive skills or are not willing to invest time to form rational beliefs. Analyzing how groups with different socio-economic characteristics vary in their subjective expectations makes it possible to test the assumption of rational expectations for the population as a whole (although it will not be possible to determine who has rational expectations and who has not, or which mechanisms drives the non-rationalities). Moreover, it is of interest to analyze to what extent reported beliefs are explained by observable respondent characteristics or contain additional information at the individual level.

Second, misguided expectations may have a negative impact on future well-being of vulnerable groups in society (see, for example, Rohwedder and van Soest, 2006). In particular, overly optimistic beliefs may lead to "under-saving."[1] It is therefore important to see if different socio-economic groups have realistic views of the future, and to what extent their misconceptions could impair their future well-being. This information could be useful for economists concerned with the mechanisms behind the formation of beliefs and could offer policy makers new ways of designing effective solutions to improper saving among the different groups.

Since the summer of 2006, monthly survey data were collected on the expectations of Dutch households concerning occupational pensions, old age social security, and the average retirement age ten or twenty years from the time of the interview. The same data (but for a shorter time period) have been analyzed by Van der Wiel (2008), analyzing the relation between these expectations and savings decisions and by Van der Wiel (2009), focusing on the effect of the number of newspaper articles on the volatility of social security expectations. We will investigate how social security, occupational pension and average retirement age expectations have changed over time and how they vary with socio-economic characteristics. Since we use data collected up to September 2010, we can also analyze the effect of the recent crisis.

The remainder of this paper is organized as follows. In Section 4.2, we describe the sample design and the expectations questions. Section 4.3 describes how the answers vary over time and associates this with the public policy debate in the Netherlands. In Section 4.4, we analyze some empirical models relating pension expectations to background characteristics. Section 4.5 concludes.

4.2 SAMPLE DESIGN AND SURVEY QUESTIONS

The survey was administered by members of the CentERpanel, an ongoing Internet panel managed by CentERdata, a data collection and applied research institute affiliated with Tilburg University. The sample is based upon a simple random sample from the population in the Netherlands of ages 16 and older and consists of over 2000 households in which one or more adults complete questionnaires at home every weekend over the Internet. Households without Internet access are provided with Internet access by CentERdata so that the survey also covers households without Internet or without a personal computer. About 75% of all panel members respond to the questions in a given weekend. Rich background information about the panel respondents is available from previous interviews.[2]

---

1 On the other hand, a recent study by De Grip, Lindeboom and Montizaan (2009) also suggests that there is a direct effect of expectations on well-being, implying that the effect of overly optimistic beliefs on life-time well-being is not unambiguously negative.

2 See http://www.centerdata.nl/en/TopMenu/Wat_doen_we/Dataverzameling/CentERpanel/index.html

Each respondent answers the questions to the specific survey on pension expectations once every three months. The total sample of respondents of ages 25 and older was randomly split into three subsamples of about the same size. One subsample gets the questions in January, April, July and October; the second subsample in February, May, August and November, etc. This implies that there are observations for one third of the sample in each month.[3] In this study, we draw on all the data collected between May 2006 and September 2010.

In addition to the questions on future expectations that we will analyze, the survey asks questions on other pension related issues, such as the respondents' satisfaction with several aspects of their pension provisions and the pension system in general; see De Bresser and van Soest (2009).

The expectations questions have been asked in the form of subjective probabilities. According to Manski (2004), this is a much better way to elicit information on people's subjective distributions of future outcomes, providing more information than, for example, simple point expectations. Subjective probability questions have been extensively used and validated in US surveys, particularly the Health and Retirement Study, which has subjective probability questions on expected retirement age, on expected old age social security income, on expected length of life, on future health problems that limit the ability to work, and on the probability to leave a bequest (see Juster and Suzman, 1995; Hurd, 2009).

The first questions are about old age social security benefit levels (AOW: Algemene Ouderdoms Wet). According to the current system, everyone who has been a resident of the Netherlands from age 15 to age 65 is fully eligible for these benefits. The amount is determined by the official minimum subsistence level[4] and depends on partnership status but usually not on earnings or employment history. There is one exception that may matter for expectations: if one spouse is older than 65 and the other is younger than 65, the couple receives the amount for singles if the younger spouse has a paid job, but the full amount for the couple if the younger spouse does not do any paid work; the additional amount received in the latter case is called the "partner allowance". It will be abolished in 2015, and this has already been announced long before the start of our survey in 2006. Respondents who are aware of this announced reform may incorporate it in their expectations concerning future benefit levels. The wording of the first series of questions was:

---

and http://www.centerdata.nl/en/TopMenu/Projecten/DNB_household_study/.

3 In May and June 2006 (the first two months of the survey) everyone was invited to participate instead of one third of the sample.

4 The 2010 amounts (including vacation allowance) are €1075 for singles and €1478 for couples.

What do you think is the probability that 10/20 years from now the purchasing power of AOW benefits will on average be

- *Less than now?*
- *At least 10 percent less than now?*
- *More than now?*
- *At least 10 percent more than now?*

Please answer on a scale from 0 to 100 percent, where 0 means it will definitely not happen and 100 means it will certainly happen.

Half of the sample got the questions with 10 years from now; the other half with 20 years from now, with randomized assignment.[5] All answers from 0 to 100 were allowed for; consistency restrictions (e.g., second answer larger than the first one) were not imposed and were indeed sometimes violated by the respondents. Note that the first and third answer may well add up to less than 100 since people may attach a positive probability to the event that purchasing power remains the same. This applies in particular to the purchasing power of AOW benefits since, in the current system, they are fixed at the minimum subsistence level and reforms proposed up to now do not change that (though for couples to whom the "Partner allowance" applies, the purchasing power of the total benefit will decrease in 2015 – see above).

The second set of questions concerns the purchasing power of second pillar pensions. Essentially all employees in the Netherlands participate in mandatory pension schemes organized at the firm or industry level, which in most cases guarantees them a defined benefit occupational pension that increases with their earnings. There are differences, however, in, e.g., how the pension level varies with the pattern of life cycle earnings or whether pension benefits keep track with inflation. The wording of the questions was similar to that for AOW benefits:

What do you think is the probability that 10/20 years from now the average purchasing power of occupational pensions will be

- *Less than now?*
- *At least 10 percent less than now?*
- *More than now?*
- *At least 10 percent more than now?*

5 This randomization was independent across waves, so the same person could get the questions with 10 years in one wave and with 20 years in another wave; in a given wave, all questions (in all four sets) for a given respondent had 10 years, or they all had 20 years.

Please answer on a scale from 0 to 100 percent, where 0 means it will definitely not happen and 100 means it will certainly happen.

The answers to these questions may be affected by the problems faced by occupational pension funds due to the financial crisis. Many pension funds have experienced a reduction of the accumulated pension wealth of their clients due to falling stock prices, and have announced that they will not compensate pension amounts for inflation in the near future in response. In the long run, this may lead to much lower pension levels in real (purchasing power) terms. Implicitly, the respondents are asked to forecast how much of the inflation in the next ten or twenty years will not be compensated by increases in nominal pensions – not an easy task.

The third set of questions is about the eligibility age for old age social security benefits:

What do you think is the probability that 10/20 years from now the age at which people are entitled to AOW benefits will on average be

- *Higher than now?*
- *At least two years higher than now?*
- *Lower than now?*
- *At least two years lower than now?*

Please answer on a scale from 0 to 100 percent, where 0 means it will definitely not happen and 100 means it will certainly happen.

This question touches the core of the Dutch policy discussion since 2008, which focuses on raising the eligibility age for AOW benefits from 65 to 66 or 67 for cohorts that will reach age 65 after a certain date (this date is also part of the discussion).[6]

The final set of questions we will analyze refers to the retirement age.[7] The wording of the questions about the retirement age is:

What do you think is the probability that 10/20 years from now the age at which people stop working will on average be

- *Higher than now?*
- *At least two years higher than now?*

6 The plan launched in September 2009 was to implement the changes 10 years from now, not affecting those who are currently older than 55; this plan was not implemented because the government stepped down, and the debate is still ongoing.

7 We did not feel it was useful to ask about the eligibility age for occupational pensions, because with increasing flexibility and actuarially fair choices, the formal eligibility age can be quite low but with unattractively low pension benefits this is not very meaningful.

- *Lower than now?*
- *At least two years lower than now?*

Please answer on a scale from 0 to 100 percent, where 0 means it will definitely not happen and 100 means it will certainly happen.

Although the current policy debate is more about postponing AOW benefits than about fixing the retirement age, the common view is that later entitlement to AOW benefits will also lead to later retirement.

4.3 TIME TRENDS AND AGE PATTERNS IN PENSION EXPECTATIONS

During the time period covered by our data, there have been several lively policy debates on public and private pension reforms. Long before the financial and economic crisis, policy makers already saw the need to reform the public pension system due to the aging of the population (see, for example, Bovenberg and Gradus, 2008). The rising government budget deficit during the crisis starting in 2008 has strengthened the need for reforms of state pensions, but, partly due to the resignation of the government in early 2010 and the long time it took to form a new government, final decisions have still not been made. Occupational pension funds, confronted with negative returns on their investments in the stock market, have emphasized the need to reduce the generosity of pension benefits, involving lower benefits or later retirement, to avoid pension premiums continuing to rise. In this section, we investigate how the general public's expectations of the generosity of the pension system have changed during the time period 2006 - 2010 and to what extent they have responded to the policy discussion.

Figure 4.1 shows how the average answers to the probability questions on the purchasing power of AOW benefits have developed over the time period covered by the survey (May 2006 - September 2010).[8] Before discussing the time patterns, some other findings are worth noting. First, the average probabilities are consistent, in the sense that the first probability ("less than now") always exceeds the second one ("at least 10% less than now"), the third probability ("more than now") always exceeds the fourth one ("at least 10% more"), and the sum of the first and third probability is always much less than 100%, implying that, on average, a substantial positive probability of about 30% is attached to the event that the purchasing power of AOW benefits will not change. This is in line with the notion that receiving AOW benefits

8 The figures are weighted with sample weights to correct for unit non-response related to gender, age, and education. We do not address the issue of possible focal answers (see, e.g., Fischhoff and Bruine de Bruin, 1999), since heaping at 0, 50 or 100 does not seem such a large problem in these data, with the percentage of 50-50 answers varying between 12 and 17 percent.

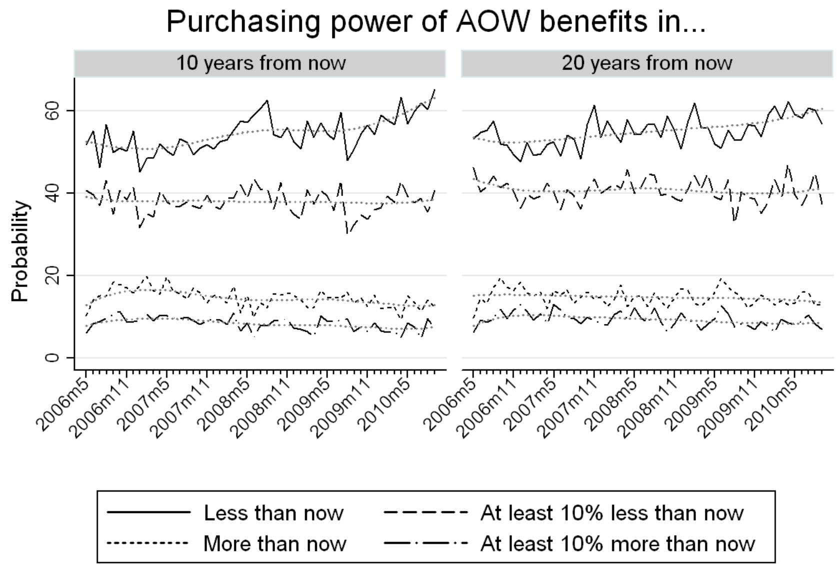


Figure 4.1: The probability of changes in the purchasing power of the AOW benefits 10 or 20 years from now

should only put household income on the official poverty line, giving a fixed purchasing power level over time.

Second, the figures are asymmetric, revealing a general sense of "pessimism": the average probability that purchasing power will fall is much larger than the probability that it will rise; and the average probability that purchasing power will fall by at least 10% is much larger than the probability that it will rise by at least 10%. This may seem surprising since there are no plans to change the purchasing power of these benefits, which, as explained above, are in principle determined by the official poverty line. On the other hand, it might reflect that some respondents are aware of the future removal of the "Partner allowance", which, although it applies to a limited subgroup of the elderly only, will reduce the average benefit per person or per household.

Third, there seem to be no systematic differences between the "10 years from now" and the "20 years from now" probabilities, although there are some non-negligible differences in specific months. Perhaps most respondents see 10 or 20 years simply as in the long run and do not make any distinction.

Hardly any time trend is found in the probabilities of an increase, an increase by 10% or more, or a fall by 10% or more. The lack of time trend may reflect the fact that the current policy debate does not concern the level of AOW benefits (the decision to remove the "Partner allowance" was already made in 1995). The only time trend we find is for the probability that benefits will fall in real terms, although even here, the pattern is not completely consistent and somewhat different

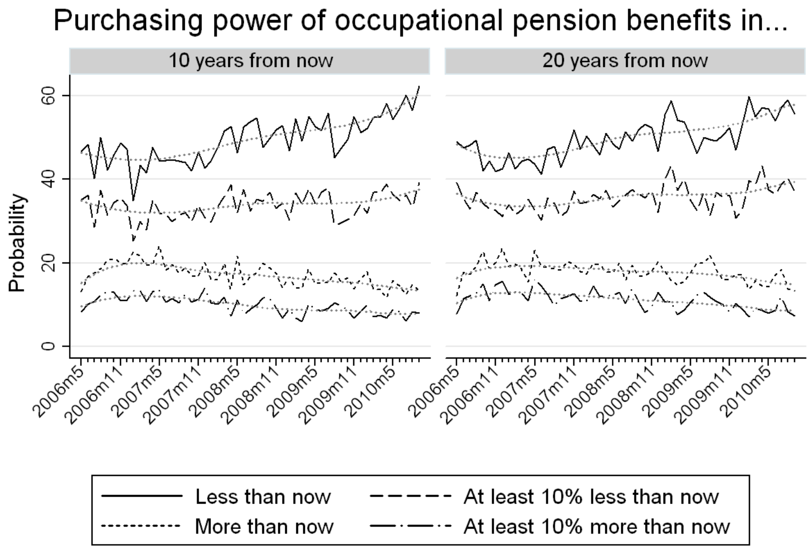


Figure 4.2: The probability of changes in the purchasing power of occupational pensions 10 or 20 years from now

for the 10 and 20 years groups. Still, we can conclude that pessimism has increased since the beginning of 2008 and particularly during the last 12 months of the survey.

Figure 4.2 shows the average answers to the probability questions on occupational pension levels, separately for the groups who got the "10 years from now" and the "20 years from now" questions. We find the same asymmetry revealing a general sense of pessimism. This is less surprising than for the state benefits, since the debate on sustainability of pensions due to the aging of the population was already quite active in 2006. Still, in principle occupational pensions are fully funded and workers save for their own occupational pension, so that population aging should not directly affect the purchasing power of these pensions if pension premiums and returns to the assets in which they are invested remain at the same level. As before, there are no systematic differences between the 10 and 20 years groups.

The trend towards larger pessimism is considerably stronger here than in the expectations concerning AOW benefits. Subjective probabilities that occupational pensions will fall in real terms have clearly risen since early 2008. This suggests that respondents have anticipated the problems that pension funds were going to face due to the financial crisis. Not much has changed in 2009, when it became clear that many pensions were no longer fully funded. The probabilities that occupational pensions will fall by at least 10% have risen as well, though by much less. Accordingly, the probabilities that the purchasing power of occupational pensions will increase or will increase by 10% or more have fallen, particularly since 2009.

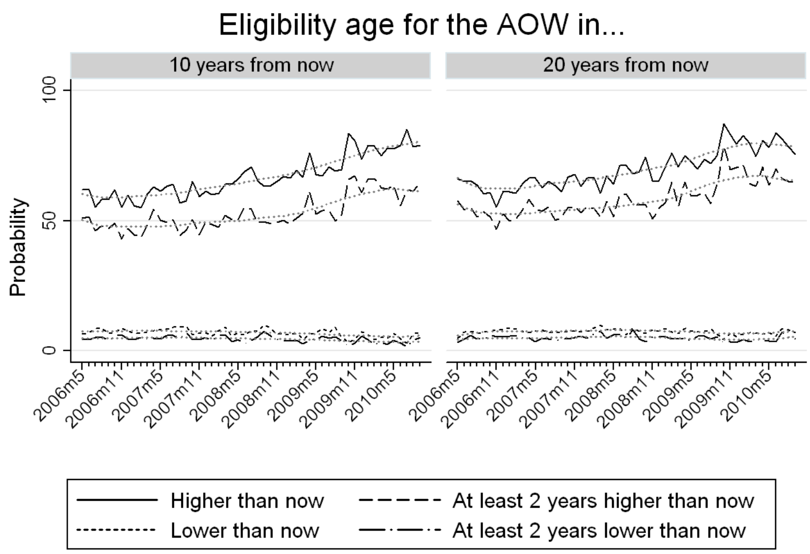


Figure 4.3: The probability of changes in eligibility age for the AOW 10 or 20 years from now

Figure 4.3 shows how expectations concerning the eligibility age for state benefits change over time. Here the asymmetry is even larger than for the pension and AOW benefit levels. The average reported probability that the eligibility age will fall (or will fall by at least 2 years) is quite small and might be upward biased by reporting errors of respondents who did not understand the questions or did not answer them seriously. The average reported probability that the eligibility age will increase over the next ten years was already about 60% in May 2006, rose to about 70% in the Summer of 2009 and to about 75% in Summer 2010. A similar clear trend towards more pessimism can be observed for the "20 years from now" group. The trend is quite plausible and in line with the announced reforms.

The figures also reveal that respondents were relatively pessimistic in the first few months of the survey (May and June 2006), probably due to the fact that the Social Democrats announced their intention to reduce the eligibility or generosity of AOW benefits to cope with the increasing costs due to population aging. In the months after that, these plans were weakened and other parties expressed disagreement, which is probably why respondents became less pessimistic over the summer of 2006. Respondents' optimism rose until the general elections in November 2006. Shortly after that, several groups revitalized the discussion on increasing the AOW eligibility age and labor force participation of older workers, and pessimism increased. Particularly since late 2008, influenced by the budget problems caused by the crisis, government plans to change the AOW eligibility age took concrete form, and increasing pessimism seems perfectly justified.

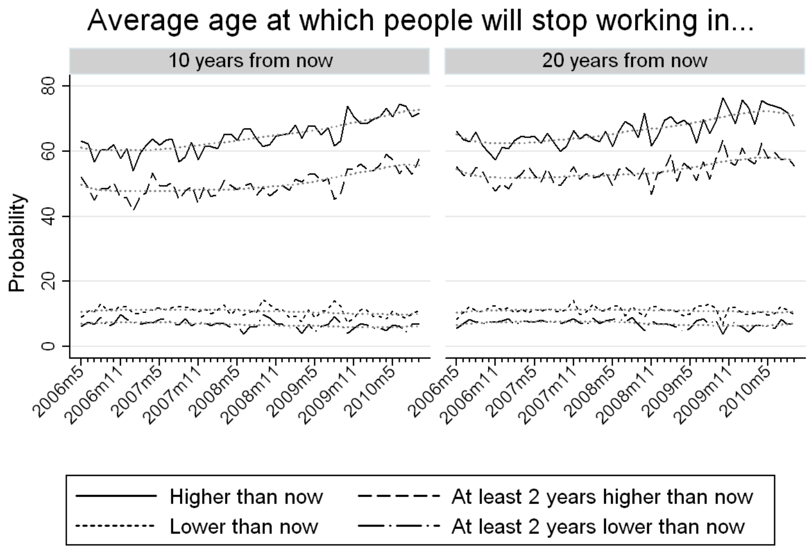


Figure 4.4: The probability of changes in the average age at which people will stop working 10 or 20 years from now

Figure 4.4 shows the development over time of expectations concerning the average age at which people will stop working 10 or 20 years from now. The asymmetry is similar to that for the AOW eligibility age. The average reported probability that the retirement age will increase over the next ten years rises from about 60% to more than 70% between 2006 and 2010. The trend is similar but somewhat less salient for the "20 years from now" group. The probability that in the next ten or twenty years the retirement age will rise by two or more years increases less, from about 50% to about 55%.

The patterns in 2006 are similar to those in Figure 4.3. People are pessimistic at first (Summer 2006) but pessimism falls until the general elections in November. In the first few months of 2007, the new government launched a plan to stimulate labor force participation of older workers by making AOW benefits dependent on participation in the years before the normal retirement age. In response to this, the number of respondents expecting an increase in the average retirement age rose. The effect disappeared when the government plans appeared to be unfeasible. In Spring 2008 the expected average retirement age rose again, possibly because some respondents already feared that the financial crisis would affect the accumulated pension wealth invested by pension funds. Respondents' expectations then remained approximately constant until the summer of 2009, but pessimism increased during the last period (Fall 2009 - Fall 2010).

The probability questions ask about general events and if everyone would have the same information set and the same way of forming their subjective distributions (like rational expectations), there should be no systematic association with respondent characteristics. We will analyze this for a large set of individual characteristics in multivariate regressions in the next section. Here we present the relation between the probabilities concerning changing the eligibility age for state benefits (see Figure 4.3) with gender (Figure 4.5) and age (Figure 4.6). Figure 4.5 shows the time pattern for men and women separately. The trend is almost identical for men and women. In most time periods, the two curves on the probabilities of postponing eligibility suggest that men are somewhat less pessimistic than women, but the differences are small.

Figure 4.6 shows how the subjective probabilities vary with respondent age, combining data from all available time periods.[9] These figures show that pessimism concerning the state pension eligibility age falls with age. For example, the average percentage probability that the state pension eligibility age will be increased is about 60% for respondents of 30 years old, but only about 40% for respondents aged 70. The average probability that the same eligibility age will rise by at least two years is about 40% for the youngest group and only 25% for the oldest age group. A surprisingly similar age pattern is found for the other questions (results available upon request from the authors) and the age patterns seem even stronger than the time trends discussed above. Interpreting the age patterns in terms of general optimism or pessimism, these results suggest the opposite of those of Dominitz and Manski (2005), who find that young people have more optimistic expectations on equity returns than older people.

4.4 Empirical models of beliefs

In this section, we will assess the impact of respondents' demographic characteristics on their reported retirement expectations. We are interested in knowing if some groups in society are particularity pessimistic or display unwarranted optimism toward retirement. As emphasized in Section 1, there are several reasons why we think this is important: to test the rational expectations hypothesis and to determine the usefulness of collecting this type of information at the micro level in future surveys, and to analyze the potential negative impact of misguided perceptions of the future on well-being for vulnerable groups in society, in particular through "under-saving."

4.4.1 *Model specification*

Following the concerns expressed above and the descriptive results in the previous section, we focus our attention on the questions concerning negative outcomes. Given the current situation of pensions,

9 Estimations obtained using local linear regression with Gaussian kernel and a bandwidth of 2 years.

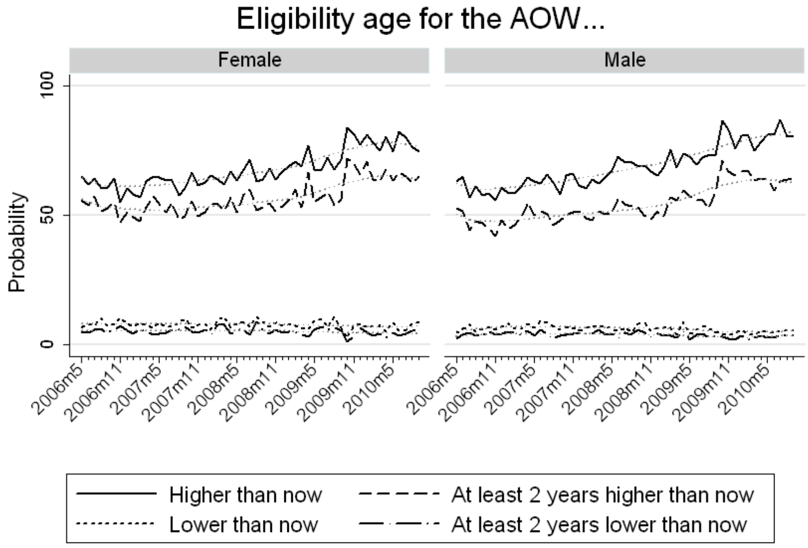


Figure 4.5: Expectations concerning the eligibility age for AOW benefits 10 or 20 years from now for men and women

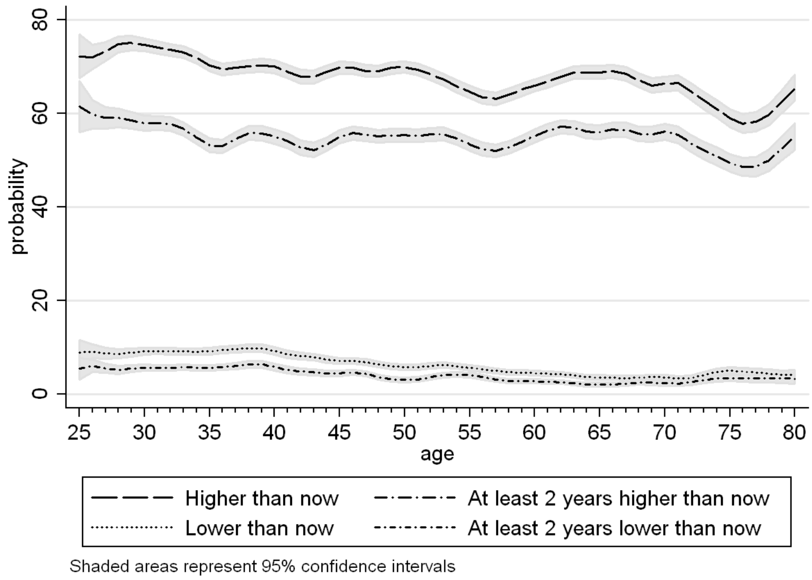


Figure 4.6: Expectations concerning the eligibility age for AOW benefits 10 or 20 years from now as a function of respondent age

changes to the actual policy that would curb the cost of the systems are more relevant than changes that would exacerbate them. We model eight dependent variables: the answers to the questions concerning a general decrease or a decrease of more than 10% in the generosity of the old age social security benefits (AOW) and of occupational pensions, and the answers to questions concerning a general increase or an increase of at least two years in the age of eligibility to AOW benefits and of the average retirement age in the Netherlands.

All dependent variables are subjective probabilities and take values between 0 %-points and 100 %-points, and a substantial number of respondents used these extreme values as answers: the percentage of zeros varies from 2.2% to 11.9%, and the fraction of 100% answers varies from 4.6% to 22.5%. We take into account the censored nature of the variables in our estimations by estimating two-limit Tobit specifications. Eight separate models are used for each of the probability questions concerning the more pessimistic outcomes (levels of state and occupational pensions lower or at least 10% lower; eligibility age for state pensions and average retirement age delayed or delayed by at least two years); the probabilities of the optimistic outcomes are always rather low and will not be analyzed further.

Over time, all respondents were asked to answer the questions up to 15 times, allowing us to control for unobserved heterogeneity at the respondent level, using panel data techniques. We therefore use random-effects Tobit models[10] with censoring both on the left at 0, and on the right at 100:[11]

$$P_{it}^* = x_{it}'\beta + \alpha_i + \epsilon_{it} \tag{4.1}$$

$$P_{it} = \begin{cases} 0 & \text{if } P_{it}^* \leq 0 \\ P_{it}^* & \text{if } 0 < P_{it}^* < 100 \\ 100 & \text{if } P_{it}^* \geq 100 \end{cases} \tag{4.2}$$

$$\alpha_i | x_{i1}, \ldots, x_{iT} \sim N(0, \sigma_\alpha^2) \tag{4.3}$$

$$\epsilon_{it} | x_{i1}, \ldots, x_{iT}, \alpha_i \sim_{iid} N(0, \sigma_\epsilon^2) \tag{4.4}$$

Here P_{it}^* is a latent variable, determined by a vector of explanatory variables x_{it}, an unobserved individual effect α_i and an idiosyncratic error term ϵ_{it}. The observed probability P_{it} is obtained from P_{it}^* through censoring at both ends, implying positive probabilities of reporting 0 and 100. The individual effects and error terms are assumed to follow normal distributions independent of the x_{it} , as in the standard random effects Tobit model. The model parameters (β, σ_α and σ_ϵ) are estimated jointly using maximum likelihood. Estimates are obtained using *Stata*.

10 Since many socioeconomic characteristics (education, gender) hardly vary over time we did not pursue using fixed effects models.

11 In some households, both spouses answered the questions. We do not account for the potential correlation between error terms of individual effects of respondents in the same household.

The same independent variables x_{it} were included in all eight models. First, we include a set of basic demographic and socio-economic respondent and household characteristics: a dummy variable with value 1 if the respondent is a male, age of the respondent, age-squared, a dummy variable taking the value 1 if the respondent lives with a partner. We control for education using dummies for intermediate education (secondary general or intermediate vocational) and high education (higher vocational or university), using low education (primary school only or lower vocational schooling) as the reference class. To capture employment status, we include dummy variables with value 1 if the respondent is retired, disabled (or partially disabled), unemployed, homemaker, and working in the public sector; the benchmark group are those who work in the private sector. We included time dummies for each month (except one) in order to control for macro-economic shocks on beliefs, like the financial and economic crisis. We also included a dummy variable that takes the value 0 if the question concerned a 10-year horizon and value 1 if the question concerned a 20-year horizon.

We controlled for respondents' income by including their log-income as an explanatory variable. Respondents who reported an income larger than €8000 per month were considered as outliers and removed from the sample. We also included in the model respondents who explicitly refused to answer the income questions or who reported a value of 0, a value often attributed to a refusal to answer. For these respondents, the value of the log-income was set to 0 and two dummy variables were included in the model: one for those who declared an income of 0 and one for unknown or undisclosed incomes.

In Table 4.1, we present the mean values of the explanatory variables in the first month (when everyone of age 25 and older was asked to participate in the survey) and in the last three months (when one third participated each month, so that the last three months cover the complete sample). The table shows that the means of most of the variables are quite stable over time. We also see that very few respondents did not report an income (none in the first month, 10 in the last three months). Not all respondents always answered the questions, due to refreshment, attrition, or temporary non-participation (e.g., holidays). About 1,300 respondents answered in the first month and in the last three months, but in total, 2,780 respondents took part in the survey over time. Average age is relatively high, because all respondents of age 25 and older are asked to answer the questions (with no upper age limit). Median net personal income (zeros excluded) rises from €1,200 in the first month to about €1,515 at the end of the survey period; there is no correction for inflation. The average education level also increases over time (low education is the reference category). The fraction of homemakers is falling over time, while the number of public sector employees is rising. A large fraction of all workers

Table 4.1: Means of the Explanatory Variables in May 2006 and July-September 2010

	May 2006	July-Sept. 2010
Male	0.496	0.489
Partner	0.768	0.771
Age	49.992	50.569
Log. net-inc. (if inc. > 0)	7.165	7.323
Inc. = 0	0.108	0.102
Unk. Inc.	0.024	0.007
Educ. Med.	0.389	0.410
Educ. High	0.234	0.252
Self-employed	0.035	0.061
Retired	0.202	0.193
Disabled	0.052	0.054
Homemaker	0.163	0.125
Unemployed	0.020	0.020
Public sector	0.212	0.245
In 20 years	0.512	0.482
N	1,309	1,121

Note: Means use respondents included in at least one of the regressions; means are weighted with sample weights based upon age, gender and education.

(almost 40%) are in the public sector, which is defined in a broad sense, including, for example, the (semi-public) health and education sectors. The dummy "In 20 years" has value 1 if the questions referred to 20 years from now and 0 otherwise; the time period in the questions was randomly drawn, independent of all other variables and with equal probabilities for "10 years" and "20 years" so that by design its *ex ante* mean should be equal to 0.5. The *ex post* mean is somewhat different, mainly due to non-response.

4.4.2 *Estimation results*

The estimation results are presented in Tables 4.2 and 4.3. Since we estimated the equations separately, we do not consider the correlations between the error terms or between the unobserved heterogeneity terms of the different equations. We consider a 5%-significance level in discussing which variables are significant and insignificant. Note that the models all explain the subjective probability of a negative outcome, so a positive sign in the estimates indicates an increase in pessimism if the independent variable increases.[12]

The results vary across the eight probabilities, but we observe some common patterns. First, males express significantly lower probabilities when it comes to the four worst-case scenarios, indicating that men are less pessimistic than women, in line with findings in the finance literature (Barber and Odean, 2001). For example, the estimated probability that the state benefit eligibility age will rise by at least two years is more than four percentage points higher for men than for women, keeping other characteristics constant. This is a much larger difference than the gender difference in Figure 5, where other characteristics were not controlled for.

Second, respondents with a partner are significantly more pessimistic in their four answers concerning age of eligibility to AOW and retirement age. A possible explanation is that couples are more concerned about retirement issues than singles and therefore pay more attention to the public debate. Another possible explanation could be that respondents with partners are often secondary earners working part-time, for whom income is not a good proxy to financial literacy or interest in financial matters (see below).

In general, high income individuals more often believe that 10 or 20 years from now, workers will retire later and the AOW eligibility age will rise. This view corresponds with the opinion of "financially literate" individuals. The dummies with value 1 if reported income is 0 or if no income is reported are significant for the four questions concerning eligibility and retirement ages. In these cases, the ln(income) variable is set to zero. Taking this into account implies that non- and zero-reporters are not very different from those with an average log income.[13]

Similarly, we find that people with medium and high education report a significantly higher probability for all the four more generic pessimistic questions. The effect is not significant for the four questions concerning a decrease of "at least 10%" or "more than two years". This finding is in line with the notion that pessimism is justified and the higher educated respondents tend to be better informed.

Many of the dummies on employment status were significant. Unemployed individuals appear to be significantly more pessimistic than private sector workers (the omitted category), giving significantly

12 The estimates of the slope coefficients cannot be interpreted as marginal effects on the expected subjective probabilities, due to the non-linearity of the model. The marginal effect of a covariate is equal to the estimated parameter times the probability of being uncensored; for the average respondent, this probability varies from 0.747 to 0.865 over the eight questions.

13 The average log income is about 7.25, so we should compare the coefficients on the dummies with 7.25 times the coefficient on log income.

Table 4.2: Estimation Results Two-Limit Tobit Models with Random Effects: Probabilities of Negative Changes in Future Generosity of State and Occupational Pensions

	Generosity of AOW		Generosity of occ. pension	
	Less...	At least 10% less...	Less...	At least 10% less...
Male	-0.864	-4.711***	-1.977	-4.596***
	(-0.679)	(-4.210)	(-1.597)	(-4.219)
Partner	2.777**	2.808***	1.849*	1.547
	(2.407)	(2.731)	(1.649)	(1.550)
Age	0.782***	0.507**	0.226	-0.009
	(3.166)	(2.317)	(0.940)	(-0.042)
Age-sqr./100	-1.198***	-0.719***	-0.673***	-0.262
	(-5.034)	(-3.405)	(-2.907)	(-1.279)
Log. net-inc.	0.819	-0.031	-0.451	-1.092
	(1.067)	(-0.044)	(-0.604)	(-1.627)
Inc. = 0	4.785	-0.610	-3.668	-6.865
	(0.919)	(-0.128)	(-0.724)	(-1.504)
Unk. Inc.	4.808	4.959	-1.683	-0.342
	(0.773)	(0.876)	(-0.279)	(-0.063)
Educ. Med.	6.369***	2.386**	3.657***	0.395
	(4.662)	(1.986)	(2.755)	(0.338)
Educ. High	11.140***	3.948***	6.512***	0.434
	(7.990)	(3.223)	(4.807)	(0.364)
Self-employed	-1.181	1.788	-2.901	-0.574
	(-0.568)	(0.963)	(-1.441)	(-0.319)
Retired	1.893	-0.270	-0.334	-1.130
	(1.232)	(-0.194)	(-0.223)	(-0.841)
Disabled	-2.442	-0.593	-2.812	-0.660
	(-1.162)	(-0.314)	(-1.372)	(-0.360)
Homemaker	-2.099	-3.265**	-3.451**	-4.846***
	(-1.168)	(-2.022)	(-1.976)	(-3.104)
Unemployed	2.292	4.630**	1.548	5.870***
	(1.000)	(2.230)	(0.701)	(2.957)
Public sector	1.891	0.880	2.026*	1.543
	(1.613)	(0.842)	(1.783)	(1.526)
In 20 years	0.278	3.031***	-0.311	1.936***
	(0.705)	(8.273)	(-0.813)	(5.517)
Constant	36.062***	37.890***	55.771***	52.991***
	(4.451)	(5.223)	(7.072)	(7.538)
Num. Ind.	3,030	3,027	3,033	3,032
Num. Obs.	25,899	25,746	26,017	25,990
ρ	0.450	0.407	0.450	0.418
σ_α	26.444	22.751	25.785	22.330
σ_ϵ	29.261	27.450	28.503	26.323

Dummies for each but the initial time period were included, but are not reported.
t-values in parentheses
Stars denote significance: * 10% level, ** 5% level, *** 1% level

Table 4.3: Estimation Results Two-Limit Tobit Models with Random Effects: Probabilities of Delays in Eligibility to AOW Benefits and Average Retirement Age.

	Eligibility to AOW		General ret. age	
	Later...	At least 2 yrs later...	Later...	At least 2 yrs later...
Male	-0.384	-5.374***	-0.644	-5.456***
	(-0.318)	(-4.500)	(-0.578)	(-5.045)
Partner	4.698***	3.610***	3.885***	3.073***
	(4.335)	(3.338)	(3.900)	(3.140)
Age	-0.588**	-0.690***	0.125	-0.029
	(-2.514)	(-2.973)	(0.583)	(-0.139)
Age-sqr./100	0.261	0.559**	-0.312	-0.029
	(1.159)	(2.497)	(-1.509)	(-0.143)
Log. net-inc.	2.547***	1.711**	2.069***	2.144***
	(3.506)	(2.343)	(3.142)	(3.272)
Inc. = 0	14.909***	11.876**	14.096***	14.721***
	(3.032)	(2.396)	(3.162)	(3.308)
Unk. Inc.	19.096***	14.266**	13.601**	18.028***
	(3.268)	(2.423)	(2.567)	(3.411)
Educ. Med.	2.638**	0.134	3.392***	0.770
	(2.040)	(0.104)	(2.846)	(0.664)
Educ. High	4.664***	-0.862	4.845***	-1.793
	(3.529)	(-0.659)	(3.977)	(-1.516)
Self-employed	-1.468	1.890	-0.946	1.624
	(-0.754)	(0.967)	(-0.531)	(0.920)
Retired	4.170***	1.866	1.016	0.808
	(2.870)	(1.286)	(0.773)	(0.619)
Disabled	-2.752	0.254	-3.671**	-0.179
	(-1.398)	(0.128)	(-2.037)	(-0.100)
Homemaker	-0.649	-2.531	-3.129**	-2.820*
	(-0.384)	(-1.493)	(-2.035)	(-1.848)
Unemployed	5.905***	6.846***	4.822**	5.913***
	(2.793)	(3.189)	(2.529)	(3.081)
Public sector	0.200	0.979	0.240	0.991
	(0.182)	(0.893)	(0.240)	(1.003)
In 20 years	4.626***	8.305***	2.775***	5.790***
	(12.664)	(22.313)	(8.445)	(17.401)
Constant	63.013***	59.007***	48.700***	39.641***
	(8.208)	(7.705)	(6.945)	(5.744)
Num. Ind.	3,033	3,032	3,035	3,035
Num Obs.	26,037	26,004	26,044	26,031
ρ	0.472	0.442	0.484	0.446
σ_α	25.208	24.738	23.542	22.532
σ_ϵ	26.667	27.812	24.325	25.110

Dummies for each but the initial time period were included, but are not reported.
t-values in parentheses
Stars denote significance: * 10% level, ** 5% level, *** 1% level

lower answers to all "worst-case" scenarios (at least 10% less, at least 2 years later) and to the questions concerning a later eligibility age for state pensions and a later average age of retirement. The coefficient on the homemaker dummy is always negative and often significant, suggesting that homemakers are less pessimistic than employees in the private sector. Retired and disabled respondents are not very different from private sector workers (everything else held constant), though retired individuals seem to believe more often that the eligibility to AOW benefits will be delayed while those receiving disability benefits less often think that the retirement age will rise. Public sector workers are more pessimistic than private sector workers but the differences are only marginally significant.

Age generally has a significant effect and the marginal effect of age is usually negative for most of the sample. The maximum of the quadratic function of age is reached between 15 and 30 years old in the series where at least one of the age parameters is significant. Note that our sample includes respondents aged 25 or older, and the large majority of the respondents are older than 30. Therefore, we can say that in general, keeping other characteristics constant, younger individuals are more pessimistic concerning the pension system than older people. This is in line with the conclusion about the age patterns in the previous section (see Figure 6), where other characteristics were not kept constant. This finding is not explained by either the knowledge or the general optimism arguments that we used above. Perhaps it relates to the fact that, in spite of the fact that the question explicitly mentions "10 years from now" or "20 years from now" respondents often answer the questions thinking about their own pension provision at the time when they retire, which will probably be less generous for younger people than for those who are already approaching retirement.

Finally, there is a significant positive effect of asking questions concerning a 20-year horizon rather than a 10-year horizon, indicating that respondents are more pessimistic concerning pension provisions 20 years from now than concerning pensions 10 years from now. This could be expected from the figures in the previous section and is in line with the fact that the effect of population aging on, for example, the ratio between the 65+ and 65- population sizes, is expected to increase further during the next twenty years.[14]

The estimates of the standard deviations at the bottom of the table (σ_α for the individual effects; σ_ϵ for the error terms) indicate that there is substantial unobserved heterogeneity: between 40 and 50 percent of the total unexplained variation in the reported probabilities can be ascribed to time persistent individual effects (as indicated by $\rho = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_\epsilon^2)$). This also answers the second question we raised at the beginning of this section: the covariates used in our model do

14 See, for example, van Duin and Garssen (2011).

not capture the heterogeneity in beliefs completely, and the reported probabilities provide additional information, (not just noise – which might be the case if σ_α were negligible compared to σ_ϵ). This is in line with the existing literature emphasizing the value of subjective probabilities in survey data (see, e.g., Manski 2004), reinforcing the idea that eliciting information of expectations is important for researchers interested in questions related to retirement and pensions.

4.5 CONCLUSION

We have analyzed expectations of the Dutch population of ages 25 and older concerning the system of income provision after retirement. The recent trends and policy discussions that seem to justify the expectation that future pensions will be less generous in terms of pension levels, eligibility ages, or both, are reflected in the trend in expectations, but only to a limited extent. Expectations seem to adjust only very slowly to the new reality and in this case this probably implies that the Dutch population is probably too optimistic, on average. Our micro-data also revealed substantial heterogeneity across and within socio-economic groups, suggesting that the average optimism is due to the over-optimism of a substantial subsample, whereas others may well have rational expectations.

The finding that men are less pessimistic than women is consistent with findings in existing studies in a different context. The fact that richer (and higher educated) individuals are significantly more pessimistic concerning some aspects of retirement than poorer respondents is in line with a positive association between socio-economic status and knowledge of the public debate on pension provisions. The finding that younger individuals are more pessimistic than older respondents may relate to the fact that respondents often answer the questions thinking about their own pension provision at the time when they retire (in spite of the wording of the questions).

From an economic policy point of view, the results we have obtained in models that relate expectations to socio-economic characteristics contain both good and bad news, under the assumption that pessimism is justified and the more pessimistic respondents are also the most realistic. That younger individuals are aware of the possible negative changes in pensions is certainly comforting news, as long as they will adapt their saving behavior accordingly. The younger individuals, who are likely to witness changes to the pension system, have time and room to adapt their employment career and their life-cycle saving plans to this new reality, and can minimize an unwanted decline of well-being at retirement.

On the other hand, we view the fact that poorer individuals tend to be more optimistic as bad news. The poorer individuals depend more on the old age social security benefits than their richer counterparts, and are therefore more affected by a reduction in the generosity of these benefits. For the poorest among them, it might not make a lot of difference to anticipate the changes, as they are not able to save for retirement and their income will probably consist almost solely of social security anyhow. However, not anticipating the policy changes could have a larger negative impact on the well-being of the middle class, who are likely to save too little under erroneous beliefs concerning the future. An unrealistic view of the future of public pensions could have important welfare effects for these respondents.

Future research opportunities remain. Adding more waves of data will help to better identify the long term consequences of the financial and economic crisis. In addition, some methodological improvements are possible. First, we already mentioned that the full information on individual behaviour provided by the multidimensional panel structure is not fully exploited. We could control for general pessimism by estimating the equations jointly, and by allowing the terms of individual heterogeneity to be correlated among individuals. Another interesting step would be to jointly analyze the beliefs of respondents within a household, and to assess if unwarranted optimism or pessimism is contagious among partners. Finally, since respondents tend to answer our probability questions using focal answers such as "50 percent," the assumptions needed for the Tobit model may not be justified, and a model that explicitly accounts for the 50-50's, other focal answers, and the rounding to multiples of 5 or 10 seems worthwhile to check the robustness of our results.

4.6 ACKNOWLEDGEMENTS

This research was funded by Stichting Instituut GAK and Netspar. We are grateful to Jan Nelissen, Rob Alessie, Viola Angelini, Mauro Mastrogiacomo, Karen van der Wiel and other participants of the conferences "Labour force participation and the well-being of the 50+ population" in Utrecht, September 2009 and "International Pension Workshop" in Zurich, June 2010.

BIBLIOGRAPHY

Barber, B.M. and Odean, T. (2001). Boys Will Be Boys: Gender, Overconfidence and Common Stock Investment, *Quarterly Journal of Economics*, 116(1), 261-292.

Benítez-Silva, H., and Dwyer, D.S. (2005). The Rationality of Retirement Expectations and the Role of New Information, *The Review of Economics and Statistics*, 87(3), 587-592.

Bovenberg, A.L., and Gradus, R.H.J.M. (2008). Dutch Policies Towards Ageing, *European View*, 7(2), 265-275.

Chan, S., and Stevens, A.H. (2008). What You Don't Know Can't Help You: Pension Knowledge and Retirement Decision-Making, *Review of Economics and Statistics*, *90*(2), 253-266.

Das, M., Dominitz, J., and van Soest, A. (1999). Comparing Predictions and Outcomes: Theory and Application to Income Changes. *Journal of the American Statistical Association*, 94, 75-85.

De Bresser, J., and van Soest, A. (2009). Satisfaction with Pension Provisions in the Netherlands: A Panel Data Analysis, Netspar Discussion Paper DP10/2009-034, Netspar, Tilburg.

Delavande, A., and Rohwedder, S. (2008). Eliciting Subjective Probabilities in Internet Surveys, *Public Opinion Quarterly*, 72(5), 886-891.

Dominitz, J. (2001). Estimation of Income Expectations Models Using Expectations and Realizations Data, *Journal of Econometrics*, 102, 165-95.

Dominitz, J., and Manski, C.F. (1997). Using Expectations Data to Study Subjective Income Expectations, *Journal of the American Statistical Association*, 92, 855-867.

——— (2005). Measuring and Interpreting Expectations of Equity Returns, NBER working paper 11313, NBER, Cambridge Mass.

——— (2006). Measuring Pension-Benefit Expectations Probabilistically, *Labour*, 20, 201-236.

De Grip, A., Lindeboom, M., and Montizaan, R. (2009). Shattered Dreams: The Effects of Changing the Pension System Late in the Game, Netspar Discussion Paper (DP 02/2009-010).

Feldstein, M. (1974). Social Security, Induced Retirement, and Aggregate Capital Accumulation *Journal of Political Economy*, 82(5), 905-926.

Fischhoff, B., and Bruine de Bruin, W. (1999). Fifty-fifty = 50 %? *Journal of Behavioral Decision Making,* 12, 149-163.

Hurd, M.D. (2009). Subjective Probabilities in Household Surveys, *Annual Review of Economics*, 1, 543-564.

Hurd, M.D., and McGarry, K. (1995). Evaluation of the Subjective Probabilities of Survival in the HRS, *Journal of Human Resources,* 30, S268-S292.

Juster, F.T., and Suzman, R. (1995). An Overview of the Health and Retirement Study, *Journal of Human Resources,* 30, S7-S56.

Keane, M., and Runkle, D. (1990). Testing the Rationality of Price Forecasters: New Evidence from Panel Data, *American Economic Review,* 80, 714-34.

Lindbeck, A., and M. Persson (2003). The Gains from Pension Reform, *Journal of Economic Literature,* 41(1), 74-112.

Manski, C. (2004). Measuring Expectations, *Econometrica,* 72, 1329-1376.

Rohwedder, S., and Van Soest, A. (2006). The Impact of Misperceptions About Social Security on Saving and Well-being, Michigan Retirement Research Center Paper (WP 2006-118).

Stephens Jr., M. (2004). Job Loss Expectations, Realizations, and Household Consumption Behavior, *The Review of Economics and Statistics,* 86, 253-269.

Van Duin, C., and Garssen, J. (2011). Bevolkingsprognose 2010-2060: sterkere vergrijzing, langere levensduur, *Bevolkingstrends, 1e kwartaal 2011*, Statistics Netherlands, The Hague, pp. 16-23.

Van Solinge, H., and K. Henkens (2007). Involuntary Retirement: The Role of Restrictive Circumstances, Timing, Social Embeddedness and Control. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences,* 62, S295 - S305.

Van der Wiel, K.M. (2008). Preparing for Policy Changes: Social Security Expectations and Pension Scheme Participation., Netspar Discussion Paper (DP 07/2008-025).

——— (2009). Have You Heard the News? How Real-life Expectations React to Publicity, Netspar Discussion Paper (DP 06/2009-031).

Zaidi, A. (2010), *Fiscal and Pension Sustainability: Present and Future Issues in EU Countries,* Policy Brief, European Centre, Vienna.

5 INDIVIDUAL SURVIVAL CURVES COMBINING SUBJECTIVE AND ACTUAL MORTALITY RISKS

This chapter is based on my joint work with Michael Hurd and Pierre-Carl Michaud.

5.1 INTRODUCTION

Mortality expectations play a key role in economic models of savings (Hurd, 1989), retirement (French, 2005) and insurance (Yaari, 1965). Most of the empirical work in these fields has used the convenient rational expectations (RE) assumption replacing individual expectations with actual life-tables aggregated to allow only for differences in gender and sometimes race. Of course, there is no particular reason why this assumption should hold, in aggregate and/or at the micro level. Some suggestive aggregate evidence against the RE has been presented by Hurd and McGarry (1993) using subjective survival probability questions from the *Health and Retirement Study* (HRS). In this paper, Hurd and McGarry showed that men were slightly more optimistic than the life-table and women slightly more pessimistic than the life-table. Using a more formal model estimating individual subjective survival curves, Gan, Hurd and McFadden (2005) showed that on average, respondents from the AHEAD cohort of HRS were optimistic relative to the life-table.

At face value, these results raise doubts about the validity of the RE assumption used in economic models involving mortality expectations when these are replaced by life-table probabilities. At the micro level, both these papers revealed a tremendous amount of heterogeneity in subjective survival probabilities. The answers co-vary with known risk factors (e.g. smoking, health problems) in the expected way. In a follow-up study using the panel dimension of the data, Hurd and McGarry (2002) showed that this heterogeneity helped predict future mortality, over and above known risk factors. It is important to point out that although using aggregate expectations is likely to be misleading at the micro level, this evidence does not necessarily imply a violation of the RE hypothesis: the heterogeneity helps predict mortality and thus individuals may have private information about their mortality risk.

There are two main reasons why life-tables may not be the right benchmark to test the RE assumption. Because there is heterogeneity in mortality risk in the population, the future age-specific mortality risk of someone alive today is unlikely to match the mortality rates of older individuals alive at the same point in time. Unless one adopts a model of period and cohort effects (e.g. Lee, 1992), using current period life tables to forecast future mortality rates will miss the very persistent but complicated march of age-specific mortality rates resulting in the observed two-digit increase in life expectancy witnessed over the last 75 years. Second, even without cohort or period effects present, aggregate life-table probabilities will exhibit a flatter age profile when there is uncontrolled unobserved heterogeneity (Vaupel, 1979). This is due to spurious negative duration dependence or dynamic selection with the fittest surviving the longest.

In this paper, we use an actual mortality model estimated from 16 years of longitudinal data for the same respondents answering expectation questions. We show how to parametrize the model to test the rational expectations hypothesis at the aggregate level but also for sub-groups of the population. The framework is flexible enough to allow for a large amount of heterogeneity and allows us to recover individual subjective survival curves which incorporate the private information held by respondents about their survival. In particular, we account for the possibility that individuals provided rounded answers which we are able to filter when making predictions for individual survival curves. We discuss how, in future work, these estimates could be used in a number of applications related to consumption & savings and population projection.

The paper is structured as follows. In Section 5.2, we describe the data and discuss the probability questions used to elicit subjective survival probabilities. In Section 5.3, we present the models of subjective survival. In Section 5.4, we present the estimation results. Section 5.5 concludes.

5.2 DATA

The data used in this paper mainly come from the *Health and Retirement Study* (HRS).[1] We use nine waves of this biennial survey, from 1992 to 2008. The sample includes respondents aged 50 and older, and their spouses. Death is recorded in exit interviews and confirmed with matches to the National Death Index (NDI). Respondents for whom the vital status is unknown are also matched to the NDI based on key characteristics such as their name, etc. Since we use covariates measured in the previous wave to predict mortality at the interview, we consider respondents from waves 1 to 8 of the survey, ranging from 1992 to 2006.

1 Note that the HRS was merged with the Asset and Health Dynamics Among the Oldest Old (AHEAD), the survey used by Gan, Hurd, and McFadden (2005), in 1998. The first waves of AHEAD were collected in 1993 and 1995. As the name implies, AHEAD surveyed only respondents aged 70 and more, although their spouses, also included in the data, were sometimes younger than 70. As a result, the sample we use here also includes respondents that are younger than those included in the reference article.

Figure 5.1 shows the comparison between period life-table and one-year mortality in three waves of the HRS. For this, we used all respondents answering in that wave and used the year of death from the HRS/NDI information to compute the fraction who are known to have survived one year. These data from the HRS are weighted using respondent-level weights for each of these three years. The period life tables are obtained from the *Human Mortality Database* (www.mortality.org).

For all years shown, there is a close correspondence between HRS survival and period life tables prior to age 75. However, HRS survival is somewhat higher at older ages in 1994 and to some extent in 2000. By 2006, the difference has vanished. This difference is likely to be due to the sampling frame the HRS used. For each cohort entering the HRS study, the non-institutionalized population is sampled. The non-institutionalized population has more favorable survival prospects, thus explaining the difference with period life-tables which record all deaths. The gap vanishes as the study progresses because HRS follows respondents who enter nursing homes. Most of those aged 75+ in 1994 comes from the 1993 AHEAD study of non-institutionalized respondents born prior to 1923, whereas those aged 75+ in 2006 comes from a mix of cohorts entering in 1992 and 1998 and are thus more representative of that population by the time they are observed in 2006 since HRS follows them into nursing homes.

This sampling difference shows how crucial it is to use the actual mortality experience of those answering the expectations questions. Figure 5.1 suggests using national mortality would lead us to underestimate the survival probability of the respondents in the HRS. Even under the assumption that these respondents have rational expectations and can report them perfectly, a comparison of their beliefs and of national mortality data would suggest erroneously that they are optimistic. In turn, this could lead to economic behavior that would appear as suboptimal or ill-advised, but that would simply be coherent with the mortality risk they really face. Given that one of the aims of the paper is to test the RE assumption, using the right benchmark is clearly important. We do not want to interpret a potentially rational deviation from nationally representative mortality as a deviation from actual mortality risk.

In order to characterize subjective survival expectations and in turn derive a test of the RE assumption, we need information on subjective survival expectations. We obtain this information from the answer to the following question:

> [Using any] number from 0 to 100 where “0” means that you think there is absolutely no chance and “100” means that you think the event is absolutely sure to happen... What do you think are chances that: You will live to at least A?

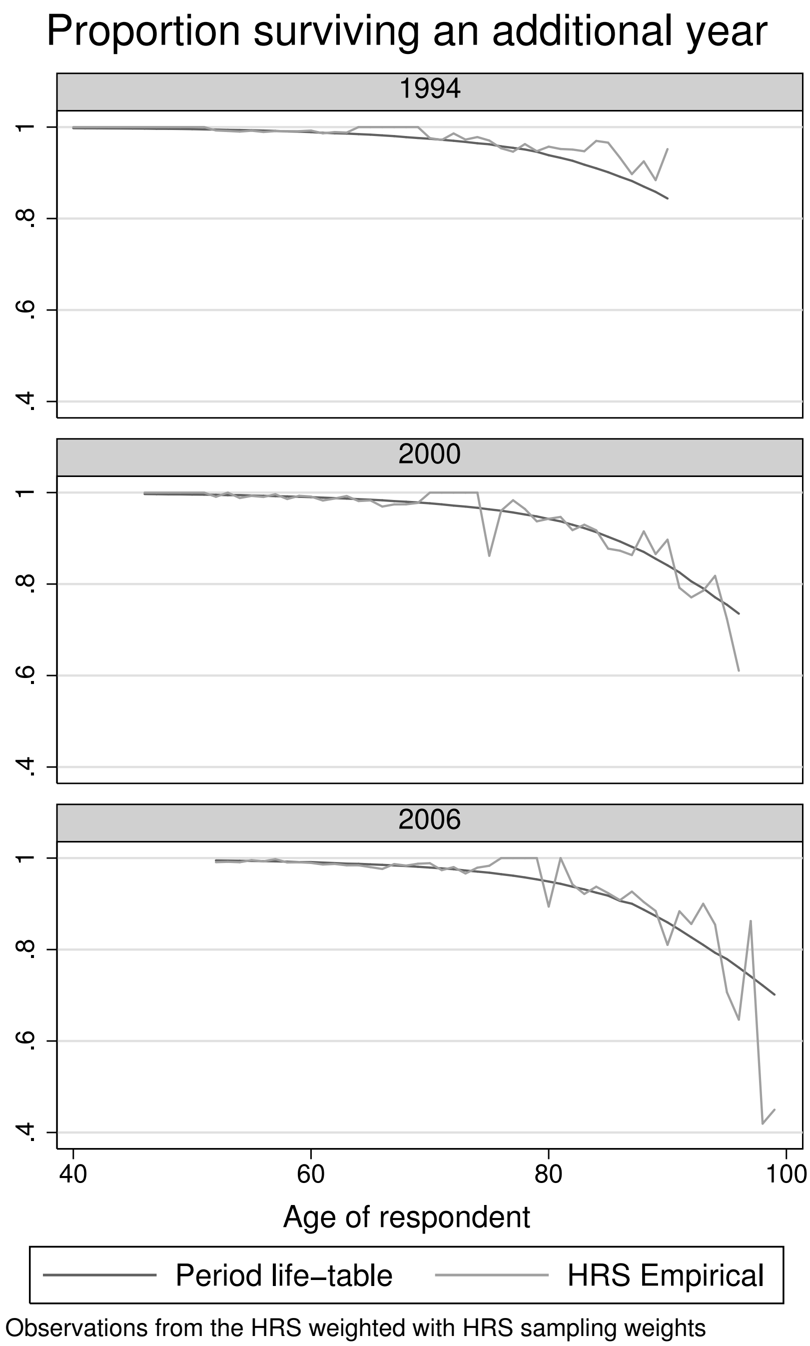


Figure 5.1: Comparison of the survival rate over one year at various ages between the HRS respondents in 1994, 2000, and 2006 with the period life-table for these years.

where A is a target age that varies for each respondent. The youngest respondents were asked to report a probability of survival to age 75. Respondents older than 65 were usually asked about survival to another target age. This target age was generally determined as an age 11 to 15 years in the future that is also a multiple of 5.[2] Our analysis includes 18,791 respondents who answered the relevant probability questions, were observed at least twice in time (first alive), and provided information on the covariates included in the analysis. When using the full sample, the number of observations (i.e. respondent-wave) is 80,298.[3]

The use of self-reported probabilities in an empirical analysis may be problematic, mostly due to rounding or measurement error. To see this, consider two series of histograms showing the distribution of the elicited subjective probabilities for the 6 target ages that were used, all waves of the HRS confounded. Figures 5.2 (a) and (b) present these histograms for females and males respectively.

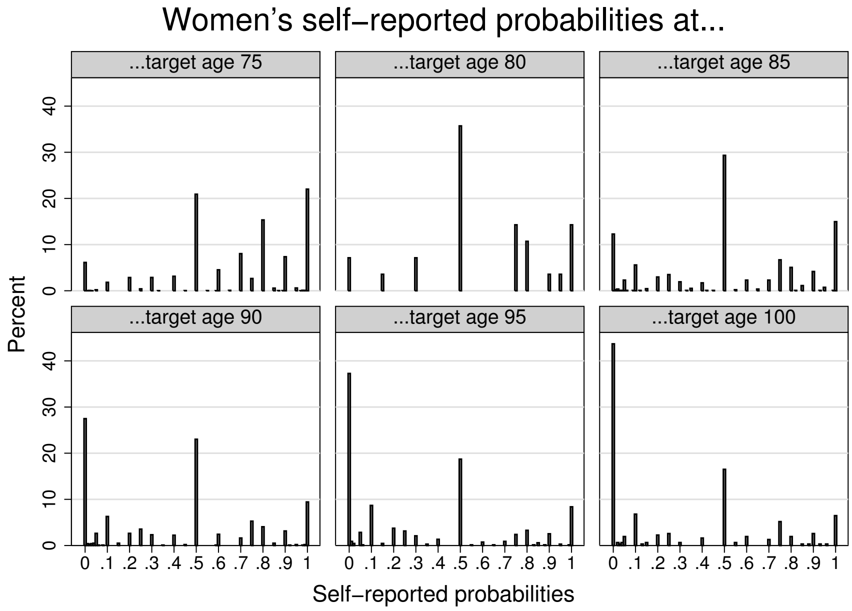
At first glance, we see that there is substantial heterogeneity in the reported probabilities, with significant heaping at the multiples of 50%-points. We can also see that the proportion of answers of that type is rather stable for each target age. Answers of 0% and 100% are particularly problematic in the analysis of survival, given that these answers are not compatible with a model of proportional hazard, often used in duration analysis. Hence, taken at face value, answers of 0% and 100% do not allow us to obtain survival curves. We also find evidence of rounding at multiples of 25, 10, and 5, and even find some very precise answers reported with a 1%-precision. The shapes of the histograms are similar among the respondents of both genders.

5.3 OBJECTIVE, SUBJECTIVE, AND SELF-REPORTED PROBABILITIES

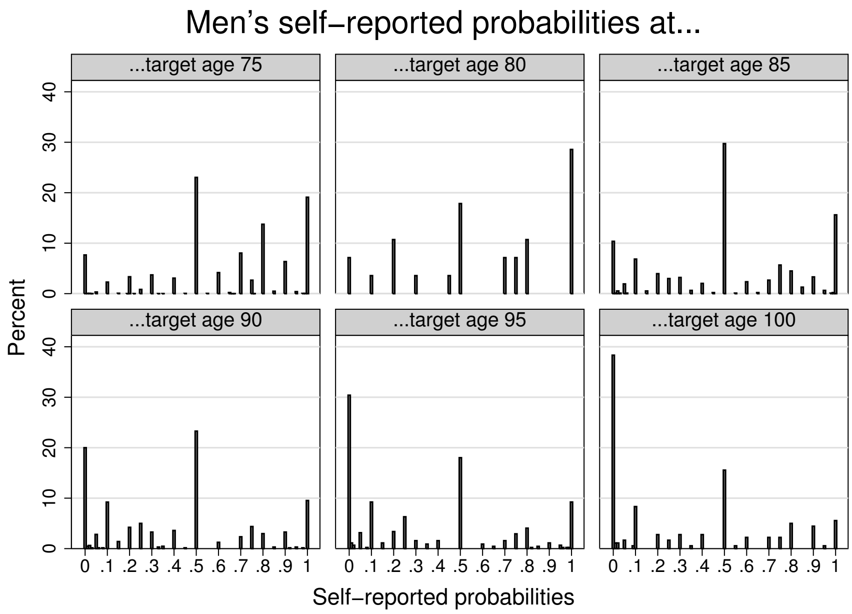
In this section, we develop a model of reporting for the subjective probabilities in the HRS. We are interested in comparing how these subjective probabilities compare to objective predictions of survival, based on within sample mortality. The objective prediction is based on an estimated model of survival, based on the literature on duration analysis. Our subjective model is parametrized in the same manner.

2 For respondents younger than 65, probability questions with a target age of 75 and 85 are asked. In what follows, we model on the response to the probability question with a target age of 75. Including the information from the second question would certainly be desirable for future work, but would require particular attention as we expect a prediction made more than 20 years in the future to be noisier than a prediction concerning a shorter horizon.

3 In this paper, we do not account for selection due to item-nonresponse to the probability questions. Since we compare actual mortality of individuals who provide a response to the subjective probability question this is less of a concern than when using life-tables. Note that respondents could explicitly answer "Don't know" to the probability question. These respondents, representing about 5% of the complete sample, were excluded from the analysis.



(a)



(b)

Figure 5.2: Histograms of the self-reported probabilities of survival to various target ages for women (a) and men (b). The target age depends on the respondent's age, hence respondents in the same sub-figures are close in age.

Thus, by analyzing the differences between the objective and subjective models, we can test the hypothesis that the expectations of the respondents can be described as rational, by testing the null hypothesis that the parameters are equal in both cases. The estimates of the objective model are obtained based on the last known survival status of the respondents, as described in Section 5.3.1. To fit a model of subjective survival, we use the probability questions discussed in the previous section. Our approach to do so is described in Section 5.3.2. However, an additional difficulty must be taken into account: the use by respondents of rounding and focal answers to express their survival probabilities, apparent in the histograms in Figure 5.2. To take this reality into account, we consider that the self-reported survival probability from current age a to target age t is a rounded value of the real underlying subjective probability. We discuss in Section 5.3.3 the model used to capture this phenomenon.

The remainder of this section introduces these three parts of our model:

1. A model for the objective hazard to predict survival among a population of respondents with given characteristics;
2. A model for the subjective hazard perceived by respondents, and for which the objective hazard is a special case;
3. A model of reporting, to take into account the rounding behavior of the respondents.

Then, Section 5.3.4 lays down explicitly the resulting likelihood function and Section 5.3.5 describes how this model can be interpreted in terms of individual subjective survival curves.

5.3.1 *Objective hazard*

We first define what the respondents should expect objectively, as a group. Our approach to measuring the objective hazard consists in estimating a survival model based on the mortality within our sample. As we already saw in Figure 5.1, the mortality within our sample is different from mortality at a national level. Thus, the use of within-sample mortality allows us to control for possible selection biases that could occur in this analysis for various reasons, such as non-response or explicit selection criteria (e.g. respondents in nursing homes could not enter the survey in the first time period). We also want to take into account that respondents entering the sample at older ages are, on average, less frail than respondents who entered at younger ages, due to dynamic selection. Additionally, the use of within-sample mortality allows us to add a socio-economic gradient in mortality risk, something that cannot be achieved with nationally representative life-tables. The fact that we allow hazard rates to vary according to observable characteristics allows us to test group-level rationality, as discussed below.

Consider the case of respondent i, aged a at the time when he answers the question concerning his probability of survival to age t. We are interested in the probability that the age of death of the respondent, denoted T, will be greater than t. This probability is given by:

$$S_a^O(t|\boldsymbol{x_i}, \nu_i^O) = \exp\left(-\nu_i^O \int_a^t \lambda^O(\tau|\boldsymbol{x_i}))d\tau\right) = \exp(-\nu_i^O \Lambda_a^O(t|\boldsymbol{x_i})) \tag{5.1}$$

where $\lambda^O(t|\boldsymbol{x_i})$ is the baseline objective hazard of respondents with characteristics $\boldsymbol{x_i}$, $\Lambda_a^O(t|\boldsymbol{x_i})$ is the integrated objective hazard from age a to age t, and ν_i^O is a frailty term added to capture unobserved heterogeneity (unobserved to the econometrician). Given that we observe respondents' survival status over a given period of time (which may include age t, but may not), the estimation of a parametric model of objective survival is straightforward and many statistical packages provide routines to do so.

We consider a continuous model with a Gompertz hazard.[4] The Gompertz hazard at age t, for a respondent with characteristics $\boldsymbol{x_i}$, constant over time, is given by:

$$\lambda^O(t|\boldsymbol{x_i}) = \exp(\boldsymbol{x_i}\boldsymbol{\zeta}^O)\exp(\alpha^O t) \qquad \alpha^0 > 0. \tag{5.2}$$

Respondents in the sample are all aged 50 or more and we do not observe death prior to that age. We therefore use age 50 as our origin (initial time at risk). The integrated hazard from the initial time period (notice that we omit the subscript a in this case) is then given by:

$$\Lambda^O(t|\boldsymbol{x_i}) = \frac{\exp(\boldsymbol{x_i}\boldsymbol{\zeta}^O)}{\alpha^O}(\exp(\alpha^O t) - 1). \tag{5.3}$$

Finally, we make the assumption that the individual frailty term ν^O is distributed with a gamma distribution with unit expectation at the origin. This parametric function depends on one parameter δ^O, and has a variance equals to $1/\delta^O$.

4 We chose the Gompertz specification as yields the best results to predict survival, based on a likelihood criterion. In order to test the robustness of our analysis to this parametric assumption, we also tried a Weibull hazard. The results we obtained were quantitatively very similar to the ones obtained with the Gompertz specification, and are therefore not reported in this paper. We chose parametric specifications of the baseline hazard over a semi-parametric piece-wise constant hazard because of the small sample of deaths at older ages (yielding high variance in the form of the hazard at older ages). The shape of the piece-wise constant hazard at younger ages was indistinguishable in terms of fit from a Gompertz specification.

Under this assumption, we can derive the expected survival probability from age a to age t as follows:

$$S_a^O(t|\boldsymbol{x_i}) = \int \frac{S^O(t|\boldsymbol{x_i}, \nu)}{S^O(a|\boldsymbol{x_i}, \nu)} f(v|T > a) dv$$

which, given that:

$$f(v|T > a) = S^O(a|\boldsymbol{x_i}, \nu) f(v) / S^O(a|\boldsymbol{x_i})$$

yields:

$$\begin{aligned} S_a^O(t|\boldsymbol{x_i}) &= \frac{1}{S^O(a|\boldsymbol{x_i})} \int S^O(t|\boldsymbol{x_i}, \nu) f(v) dv \\ &= \frac{S^O(t|\boldsymbol{x_i})}{S^O(a|\boldsymbol{x_i})} \\ &= \left(\frac{\delta^O + \Lambda^O(a|\boldsymbol{x_i})}{\delta^O + \Lambda^O(t|\boldsymbol{x_i})} \right)^{\delta^O} . \end{aligned} \tag{5.4}$$

We estimate this objective likelihood using the longest duration available for the 18,791 respondents. Each respondent is observed last at a given age, either the age at death or the age at the last interview when the respondent is still alive. We denote this age by t_i^O. We observe respondent i entering the survey at age a_i, leading to a left-censoring problem. Additionally, as some respondents are still alive at time t_i^O, we must also consider right censoring. Therefore, the likelihood contribution of an individual i is given by:

$$L^O(\zeta^O, \alpha^O, \delta^O | \boldsymbol{x_i}, a_i, t_i^O) = \lambda^O(t_i^O|\boldsymbol{x_i})^{d_i} S_a^O(t_i^O|\boldsymbol{x_i}) \tag{5.5}$$

where d_i is a dummy variable taking the value 1 if the respondent is deceased at time t_i^O.

5.3.2 *Subjective hazard*

The parametric assumptions used in the subjective hazard model are the same as in the objective hazard case. We use the superscript S to denote the subjective parameters. The use of the same parametric specification gives us a direct test of rationality. Should the agents be rational, their subjective and objective hazards would be equal. Thus, using the same parametric specification gives us a direct test of rationality by testing the difference between the objective and subjective parameters. Using a similar notation for subjective probability of survival, we define:[5]

$$S_a^S(t|\boldsymbol{x_i}, \nu_i^S) = \exp\left(-\nu_i^S \int_a^t \lambda^S(\tau|\boldsymbol{x_i}))d\tau\right) = \exp(-\nu_i^S \Lambda_a^S(t|\boldsymbol{x_i})) \quad (5.6)$$

While estimating the model, however, we are interested in the differences between the objective and subjective parameters. In order to make interpretation and testing easier, we redefine the problem in the following way:

$$\alpha^S = \alpha^O + \psi_\alpha \quad (5.7)$$
$$\boldsymbol{\zeta}^S = \boldsymbol{\zeta}^O + \boldsymbol{\psi} \quad (5.8)$$
$$\delta^S = \delta^O + \psi_\delta. \quad (5.9)$$

Hence, we define our test of rational expectations as testing the null hypothesis that all elements of $\boldsymbol{\psi}$, ψ_α, and ψ_δ are equal to 0.

For reasons that will become clear below, we also need to derive the distribution of the subjective expectations conditional on being alive at time a. Let us denote $F_s(s^S|T > a, \boldsymbol{x_i})$, the distribution function of the subjective survival probability conditional on surviving to age a. Given our previous assumptions, it is easier to recast the problem in term of $F_{\nu s}(v|T > a, \boldsymbol{x_i})$, the distribution of ν^S conditional on surviving to time a:

$$F_s(s^S|T > a, \boldsymbol{x_i}) = 1 - F_{\nu s}\left(-\frac{\ln s^S}{\Lambda_a^S(t|\boldsymbol{x_i})}\middle| T > a, \boldsymbol{x_i}\right)$$

To simplify the notation, let us denote $d = \delta^{-1}$ and emphasize that ν^S at the origin follows a distribution *Gamma* (d, δ), which given the value of d, implies a gamma distribution with unit expectation. The distribution of ν^S conditional on being alive at age a can therefore be derived in the following manner:

5 We assume that ν^O and ν^S are independent, omitting the fact that respondents with lower objective survival probabilities due to unobserved heterogeneity are also more likely to report lower subjective probabilities. Given the complexity of taking this relation explicitly into account, we leave this for future research.

$$
\begin{aligned}
F_{\nu s}(\nu|T>a,\boldsymbol{x_i}) &= \frac{1}{S^S(a|\boldsymbol{x_i})}\int_0^{\nu} S^S(a|\boldsymbol{x_i},u)f(u)\mathrm{d}u \\
&= \frac{1}{S^S(a|\boldsymbol{x_i})}\int_0^{\nu} \exp(-u\Lambda^S(a|\boldsymbol{x_i}))\frac{\delta^d u^{d-1}\exp(-\delta u)}{\Gamma(d)}\mathrm{d}u \\
&= \frac{\delta^d}{S^S(a|\boldsymbol{x_i})}\int_0^{\nu} u^{d-1}\frac{\exp(-u(\Lambda^S(a|\boldsymbol{x_i})+\delta))}{\Gamma(d)}\mathrm{d}u \\
&= \frac{1}{S^S(a|\boldsymbol{x_i})}\frac{\delta^d}{(\Lambda^S(a|\boldsymbol{x_i})+\delta)^d} \\
&\quad\times\int_0^{\nu}\frac{(\Lambda^S(a|\boldsymbol{x_i})+\delta)^d u^{d-1}\exp(-u(\Lambda^S(a|\boldsymbol{x_i})+\delta))}{\Gamma(d)}\mathrm{d}u \\
&= \int_0^{\nu}\frac{(\Lambda^S(a|\boldsymbol{x_i})+\delta)^d u^{d-1}\exp(-u(\Lambda^S(a|\boldsymbol{x_i})+\delta))}{\Gamma(d)}\mathrm{d}u
\end{aligned}
$$

from which it directly follows that $F_{\nu s}(v|T>a,\boldsymbol{x_i})$ is a distribution *Gamma* $\left(\delta^{-1},\Lambda^S(a|\boldsymbol{x_i})+\delta\right)$.

Hence, with this specification, both the expected value of the frailty term and its variance decrease with age, due to the increasing integrated hazard. This is due to the dynamic selection, where frailer respondents are more likely to die at younger age than healthier individuals. As individuals age, those who survive are more likely to be among those with a low frailty, explaining both the decrease in conditional expected value and conditional variance of the frailty term compared to the initial period.

5.3.3 *Rounding and self-reported probabilities*

Given the large number of self-reported probabilities with values of 0, 0.5, or 1 observed in the sample, it seems important to correct for the reporting behavior of the respondents. This was already acknowledged in the GHM model, relying on a censored normal distribution to update the prior distribution. This distribution, however, did not correct for the reported probabilities of 50%.

We assume that the self-reported probability p_{iat} is a rounded report of $s_a^S(t|\boldsymbol{x_i},\nu_i^S)$. We use a reporting model of the latent-variable type, comparable to the one presented by Heijtan and Rubin (1990,1991) or Kleinjans and van Soest (2010). We do not know the rounding rule used by the respondents. The problem of unknown rounding is quite simple in nature, and is comparable to a problem of unknown mixture of distributions. Suppose that we observe a respondent answering 25% to a probability question. The answer could be the result of rounding of a subjective probability in the [12.5, 37.5) range, rounding at the nearest multiple of 25. It could also come from the [22.5, 27.5) range if a respondent rounded to the nearest multiple of 5, or even from the [24.5,25.5) range should a respondent give a very precise answer. Based on distributional assumptions, we can estimate a model predicting the probability to use various rules. We consider the following rules:

1. Throw-away 50%-points, where respondents use 50%-points in order to avoid answering;[6]

2. Rounding to a multiple of 50%-points;

3. Rounding to a multiple of 25%-points;

4. Rounding to a multiple of 10%-points;

5. Rounding to a multiple of 5%-points;

6. Precise answers rounded at 1%-points.

We treat the rounding rule as an unknown random variable R. We denote the realization of R with r, an integer from 1 to 6 according to the above list. Hence, a higher value means a more precise answer or less rounding. Each rounding rule leads to a set of admissible p_{iat}. This set of admissible values is denoted by Ω_r. Finally, for each rounding rule, a self-reported probability of p_{iat} can result from rounding of values between $l_r(p_{iat})$ and $u_r(p_{iat})$. In cases where rounding is made with equally spaced intervals, we would have $l_r(p_{iat}) = p_{iat} - e_r$ and $u_r(p_{iat}) = p_{iat} + e_r$, with e_r being one half of the rounding scale.

We are interested in estimating the probability of observing a self-reported answer of p_{iat} given a subjective survival probability. The rounding process we described can be summarized as follows:

$$\begin{aligned} \text{Prob}(p_{iat}|s_a^S(t|\boldsymbol{x_i}, \nu_i^S), z_i) = \sum_{r=1}^{6} 1(p_{iat} \in \Omega_r)\text{Prob}(R_i = r, |z_i) \\ \times 1(l_r(p_{iat}) \leq s_a^S(t|\boldsymbol{x_i}, \nu_i^S) < u_r(p_{iat})). \end{aligned} \tag{5.10}$$

In order to estimate the probability of using one of these six rounding rules, we follow an approach similar to the one discussed by Kleinjans and van Soest (2010). We consider that each respondent has a "propensity to round", which is represented by r_i^*. A higher value for this variable implies that respondents are more likely to use a precise rounding rule. To capture the propensity to round, we use an ordered model. Specifically, we assume that:

$$r_i^* = z_i\gamma + \varepsilon_i$$

6 While this rule is not rounding, strictly speaking, we treat this as any answer from the interval [0,100] that is reported at the middle point. In this regard, we depart from Kleinjans and van Soest (2010) who specifically modeled the probability of giving a throw-away answer in a first step, and then modeled rounding conditional on giving a meaningful answer. Our approach differs by assuming that throw-away 50s are very imprecise answers, not non-responses.

and that a respondent uses rounding regime r if $m_{r-1} < r_i^* \leq m_r$, where where $m_0 = -\infty$, $m_1 = 0$, and $m_6 = \infty$. It follows that m_2 to m_5 are parameters to be estimated. We assume that ε_i follows a standard normal distribution, leading to a model that is analogous to an ordered probit. Equation 11 can be rewritten as:

$$\begin{aligned}\text{Prob}(p_{iat}|s_a^S(t|\boldsymbol{x_i},\nu_i^S), z_i) = \\ &\sum_{r=1}^{6} 1(p_{iat} \in \Omega_r)\,(\Phi(m_r - \boldsymbol{z_i}\gamma) - \Phi(m_{r-1} - \boldsymbol{z_i}\gamma)) \\ &\times 1(l_r(p_{riat}) \leq s_a^S(t|\boldsymbol{x_i}, \nu_i^S) < u_r(p_{iat})).\end{aligned} \tag{5.11}$$

5.3.4 *Likelihood*

We estimate the model by maximum likelihood. The approach we take is to maximize the probability of observing the self-reported answers given by the respondents. Hence, the subjective likelihood is given by:

$$L^S(\zeta^O, \alpha^O, \delta^O, \psi, \psi_\alpha, \psi_\delta, \gamma|\boldsymbol{x_i}, a_i, z_i) = \Pr(p_{iat}|\boldsymbol{x_i}, z_i)$$

which, based on the assumption made in the previous sections, is given by:

$$\begin{aligned}\text{Prob}(p_{iat}|\boldsymbol{x_i}, z_i) = \\ &\sum_{r=1}^{6} 1(p_{iat} \in \Omega_r)\,(\Phi(m_r - \boldsymbol{z_i}\gamma) - \Phi(m_{r-1} - \boldsymbol{z_i}\gamma)) \\ &\times (F_S(u_r(p_{iat})|\boldsymbol{x_i}, T > a) - F_S(l_r(p_{iat})|\boldsymbol{x_i}, T > a)).\end{aligned} \tag{5.12}$$

Finally, remember that the objective parameters appear both in the objective and subjective part, and that we want to test the difference between the objective and subjective parameters. The simplest way to test this difference is to jointly estimate the objective and subjective models. The complete likelihood is given by:

$$\ln L = \ln L^O + \ln L^S. \tag{5.13}$$

This procedure does not change the estimation of the objective parameters, as objective estimates are fixed by the objective part, and the subjective parameters adjust to these values. This approach allows us to correct for the fact that the objective parameters are estimated, and that a two-step procedure where the objective parameters are estimated first and then plugged into the subjective part would underestimate the standard errors of the subjective parameters' estimates.

5.3.5 *Subjective frailty and individual curves*

One of the objectives of this paper is to generate individual subjective survival curves for the respondents. Assuming that we knew the value of ν_i^S, we could simply use the model estimated above to obtain full survival curves with our estimated Gompertz hazard. However, the model predicts a distribution of ν_i^S for a given p_{iat} rather than a precise value for this term. In the remainder of the text, we focus on the expected value of ν_i^S, conditional on the answer to the self-reported survival probability. This expectation is given by:

$$E(\nu_i^S|p_{iat},\boldsymbol{x_i},z_i) = \frac{\int_0^\infty u\text{Prob}(p_{iat}|u,\boldsymbol{x_i},z_i)f_{\nu S}(u|\boldsymbol{x_i},T>a)\mathrm{d}u}{\text{Prob}(p_{iat}|\boldsymbol{x_i},z_i)}. \quad (5.14)$$

Focusing on the expected value of ν_i^S ignores an important piece of information about the distribution of this term. In cases where a respondent gave a very imprecise answer to the probability question, the distribution of ν_i^S has a larger variance than in the case where the respondent answered precisely. Nevertheless, the results presented in the following section are based on the expected value of this term of subjective frailty. A more complete analysis taking into account the full distribution of ν_i^S is left for future work.

5.4 ESTIMATION RESULTS

We present in this section the results of the estimation of the model introduced above. We first discuss the results for the objective part of the model, and then discuss the subjective part. The last parts of the section describe the implications of our results in terms of life-expectancy, with particular attention to a group of respondents for whom life-tables are not specifically available: the respondents who have ever smoked.

5.4.1 *Objective survival curves*

We first discuss the objective model, used as a benchmark of rational expectations for the respondents. Given that we focus on respondents aged 50 or more, we use this age as the initial time at risk, and hence estimate a model of survival from age $a - 50$ to age $t - 50$. The covariates that we include, and that we use in all models through the paper, are dummies taking a value of 1 if the respondent is, respectively, male, black, or Hispanic. We then include dummies for education, indicating if the highest educational achievement is a high school degree or a college degree ("no degree" used as a reference). Additionally, we include a dummy if the respondent has ever smoked, and two cohort dummies if the respondent was born before 1930 or after 1940 (being born between 1930 and 1939 is the reference group). The selected covariates are variables that are known to influence mortality and that are generally constant over time for a given respondent. An increase in the parameters associated with these variables implies an increase in hazard, and hence, a lower survival probability. Cohort dummies are included in order to control for variations in survival probabilities for older respondents that would not be captured by the observables characteristic.

The first column of Table 5.1 presents the result of the estimation of the objective model. At first glance, the demographic characteristics all have the expected sign, although Hispanic is not significant. An interesting result is the fact that the predicted hazard increases for younger cohorts. This would suggest a decrease in life-expectancy in years to come, which is in line with Olshansky et al. (2005) but in contradiction with what Manton et al. (2006) predict. However, this effect could also be due to a selection problem. For instance, we previously mentioned that the HRS samples from the non-institutionalized population. Hence, a frail respondent aged 75 is more likely to be sampled at that age than he is at later age. This would lead to a seemingly healthier population from the older cohort, which is what we find. Generally, if frailer and older respondents decline to join the survey, we would find a hazard increasing with the more recent cohorts.

In Figure 5.3, we present individual curves for the respondents born in 1940 or later and aged 50 to 53 –the youngest respondents of our sample. We pooled these four ages together in order to have a reasonable sample size. Curves are presented by gender. To illustrate the heterogeneity in objective survival probabilities, each sub-graph presents 3 curves: the median, 10^{th} percentile and 90^{th} percentile. For the sake of comparison, we also include the period life-table survival curve for the year the median respondent answered the survey. Given the fact that life-tables concern the complete American population, the percentiles for this figure are based on weighted estimates in HRS. Subsequent figures give equal weight to each respondent.

While the median curve for the males seems relatively close to the life-table prediction, our objective model predicts higher mortality than expected for women. This finding was also observed by Hurd and McGarry (2002) who reported that "Whereas women in the HRS appear to underestimate their survival probability, they also died at greater than expected rates (p. 983)."[7] This finding highlights one of the main advantages of the approach we propose here: even with a nationally representative sample like the HRS, the life-tables may not be a valid measure of within-sample survival probabilities. Deviation

7 We also performed sensitivity analysis using alternative specifications and estimating the model separately by gender, and the mortality among women still remains higher than predicted. Results obtained with these alternative specifications were hardly different from the ones presented here.

Table 5.1: Estimation results

	Objective model[a] (ζ^O and α^O)	Subjective model[b] (ψ, ψ_α, and ψ_δ)	Reporting model (γ and m_r)
Male	0.2864	-0.1642	0.0117
	(0.0305)	(0.0322)	(0.0104)
Black	0.2367	-0.2399	-0.1021
	(0.0413)	(0.0438)	(0.0150)
Hispanic	-0.0943	0.4094	0.0018
	(0.0604)	(0.0641)	(0.0206)
High School	-0.2173	-0.0942	-0.0090
	(0.0348)	(0.0375)	(0.0135)
College	-0.3982	-0.1964	0.1853
	(0.0361)	(0.0386)	(0.0135)
Ever Smoked	0.4572	-0.2755	0.0120
	(0.0342)	(0.0359)	(0.0109)
Cohort Prior to 1930	-0.4508	0.2554	0.2132
	(0.0570)	(0.0639)	(0.0128)
Cohort 1940+	0.2123	-0.2863	0.0141
	(0.0574)	(0.0584)	(0.0121)
α	0.1126	0.0028	
	(0.0039)	(0.0041)	
δ^{-1}	0.0815	1.0124	
	(0.0416)	(0.0421)	
Immediate word recall (z-score)			0.0169
			(0.0066)
Prop. of 0/50/100 in other quest.			-1.285
			(0.0212)
Constant (ζ_0, ψ_0, or γ_0)	-6.0868	0.9598	4.5601
	(0.0836)	(0.0879)	(0.2448)
m_2			3.6333
m_3			4.0063
m_4			5.2925
m_5			6.3240
N	18,791	80,298	
Log-likelihood		-219,682.605	

Standard errors in parentheses

[a] Survival model estimated using age 50 as initial time period.

[b] Subj. parameters expressed in terms of differences as in equations 5.7 to 5.9.

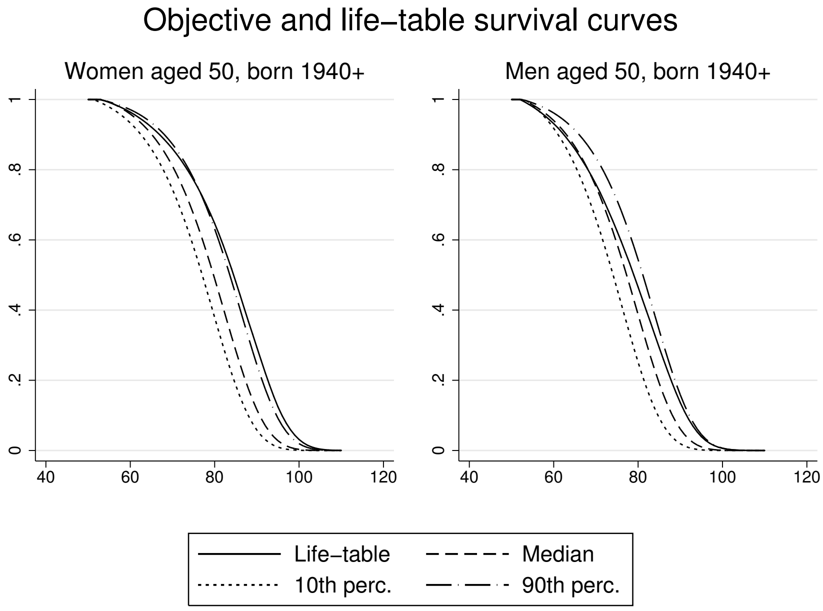


Figure 5.3: Comparison of the objective and of the life-table-based survival curves. Percentiles in this figure were determined using HRS weights.

from the life-tables in subjective expectations may not be due to an erroneous prediction.

5.4.2 *Subjective survival and reporting*

We now discuss the core result of this paper: the estimation of the full model of subjective expectations. The second column of Table 5.1 contains the main parameters of interest: the estimated subjective parameters $\boldsymbol{\psi}$, ψ_α, and ψ_δ, capturing the differences between the objective and subjective parameters. Remember that we are interested in deviations from the value 0. While it is difficult to assess the importance of the values of the parameters directly from their values, we will present below some predicted subjective survival curves resulting from our model, making it easier to picture the consequences of these variations in subjective hazards. Nevertheless, here are some preliminary remarks.

First, notice that the variance of the frailty term is much larger for the subjective frailty than it was for the objective frailty, meaning that there seems to be higher perceived variation in health from the respondents than what we found with objective mortality. Taking into account the fact that the difference between α^O and α^S is quite small and does not appear to be significant (p-value = 0.497), and taking into account that the difference in constant terms is larger than 0, it appears that the subjective curves, unconditional on the individual frailty, are more pessimistic than the objective curves at younger ages. However, due to dynamic selection, the unconditional curves tend to be more and more optimistic as time goes by, eventually leading to optimism at older ages.

Let us now consider the additional parameters to the model. The covariates' parameters can be seen as an additional correction toward optimism or pessimism that would affect all time periods. With the dynamic selection, however, each parameter individually cannot be seen as a shift away from rational expectations. Still, we see that men, black respondents, educated respondents (either with a high school or college degree), respondents from the younger cohorts, and smokers have a negative correction in their hazard. Only Hispanic respondents have a significant positive correction in their hazard, leading to lower subjective survival rates than the other respondents, everything else being held constant. It is again hard to interpret directly the meaning of these parameters. Section 5.4.4 provides an analysis for the case of smokers, illustrating the implication of our results.

We proposed above, as a test of rational expectations, to test whether all the parameters capturing the differences between the objective and subjective parameters are equal to 0. A Wald-test of joint significance rejects strongly the null hypothesis of rationality at any meaningful significance level. This result is not surprising given the difference in variance between the objective and subjective frailty.

The results for the rounding model are presented in the last column. We include two additional explanatory variables in the reporting model. The first one is based on the number of words that a respondent could recall in a memory exercise, in order to capture cognitive skills. The variable used is a period z-score, to correct for the varying number of words asked in total and for possible variation in the difficulty of the list in a given period. The second variable is the proportion of answers that were multiples of 50 in the other probability questions of the HRS. We find that the variables male, Hispanic, high school, and "ever smoked" do not have a significant effect on rounding behavior, and find that college education and cognitive skills lead to respondents giving more accurate answers. We also find that a higher proportion of focal answers in other HRS questions leads to a higher probability of coarse rounding in the self-reported survival probability, reinforcing the idea that some respondents are simply less prone to give "precise" answers. Based on our results, we do not find any support for the idea that 50%-point answers are used to avoid answering questions. In our model, the predicted probability of having this behavior is practically nil for all respondents. On average, we predict that 33.7% of the respondents round to the nearest multiple of 50, while 15.8% to a multiple of 25, 37.8% to a multiple of 10, 11.0% to a multiple of 5, and 1.7% are expected to report very precise answers.

Is rounding an important part of our model? To answer this question, we estimated a model where we fixed the propensity to report precise answers at an arbitrarily large value, implying that respondents reported their subjective probabilities with a 1%-point precision. Results are not reported here but are available upon request. Among the changes in the estimated parameters, one is very important: the variance of the subjective frailty term is about twice as large when we do not take rounding into account, increasing to a value of 2.56 with a standard error of 0.04. Other parameters seem to adjust to this change. Given the histograms presented in Figure 5.2, this is not surprising. The variation in frailty needed to accommodate the large fraction of respondents who used 0% and 100% as answers must be quite substantial. If we believe that respondents used rounding to report their probability, which we think is fairly uncontroversial based on the shape of the histograms, then not correcting for this rounding would lead us to overstate the optimism of optimistic respondents and the pessimism of the pessimistic respondents. We therefore believe that it is better to correct for rounded answers.

As previously mentioned, it is hard to assess the implications of the estimated parameters in Table 5.1. To illustrate graphically some of our results, Figure 5.4 presents some of the subjective survival curves resulting from the estimation of the full model. They also include the objective survival curve of the median respondent. On the left, the two sub-figures are not conditional on p_{iat}, the self-reported probabilities. In this case, we do not have information on the subjective frailty term, and we must therefore compute the curves based on Equation 5.4. Note that the difference in these curves is only due to variation in observable characteristics. The conditional curves on the right includes the information contained in p_{iat}. For each respondent, we compute the expected subjective frailty term as described in Equation 5.14 of Section 5.3.5. The survival curves based on these values are computed. Therefore, part of the variation in these curves is due to variation in subjective frailty and part is due to differences in observable characteristics.

Looking at the unconditional curves shows that observable characteristics do not seem to explain much in terms of variation in subjective probabilities. We see evidence that women appear to be relatively rational in their expectations concerning mortality, if we consider the median curve. As expected, the unconditional curve is flatter than the objective curve. When looking at the conditional curves, we see that the median objective and subjective curves are quite close. This would imply that women are conscious of the higher-than-predicted mortality risk. We also see evidence that men are optimistic. The unconditional curves show evidence of lower expected mortality at higher age, as we expected from Table 1. We see that the median conditional subjective curve is close to the median objective one. However,

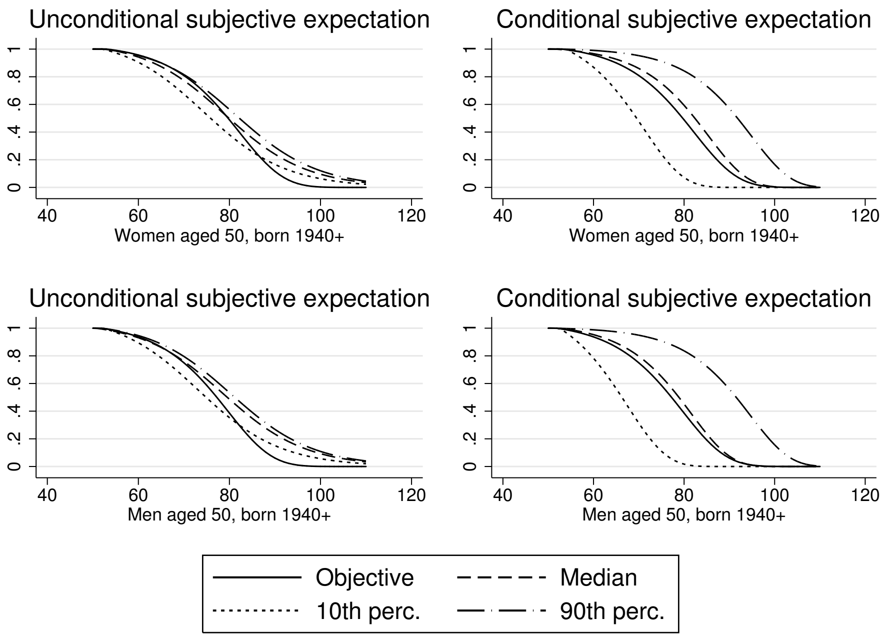


Figure 5.4: Subjective survival curves unconditional and conditional on the respondents' self-reported probabilities (expected subjective frailty used).

based on the shape of the unconditional curve, we can deduce that it implies a strong asymmetry in frailty, hinting that a lot of men are very optimistic and have curves close to the represented 90th percentile. We also see from the dispersion of the curves that while the observable characteristics account for only a small fraction of the heterogeneity, conditioning on the self-reported probability increases considerably the dispersion between the 10th and 90th percentile.

5.4.3 *Subjective life-expectancy*

For the study of pension and aging, the probability of survival at target age tells only part of the relevant story. We are not only interested in describing the perceived probability to be alive at a target age, but are interested in determining, for instance, the number of years a respondent expects to survive beyond retirement age.

Let us first consider the life-expectancy of the respondents aged 50 to 53 in our sample. Figure 5.5 presents the predicted age of mortality by gender both unconditional (first column) and conditional on probability answers (second column). As expected, considerable variation comes from the answers to the probability questions. We can see that the conditional histograms have peaks at some values, driven by the fact that respondents tend to use rounded values in their answers. Remember that we focus here only on the expected values. It follows that respondents with similar characteristics and reporting the same probability of survival have the same expected subjective frailty, leading to the same subjective life-expectancy. In fact, the predicted life-expectancy of each respondent is better described by a distribution.

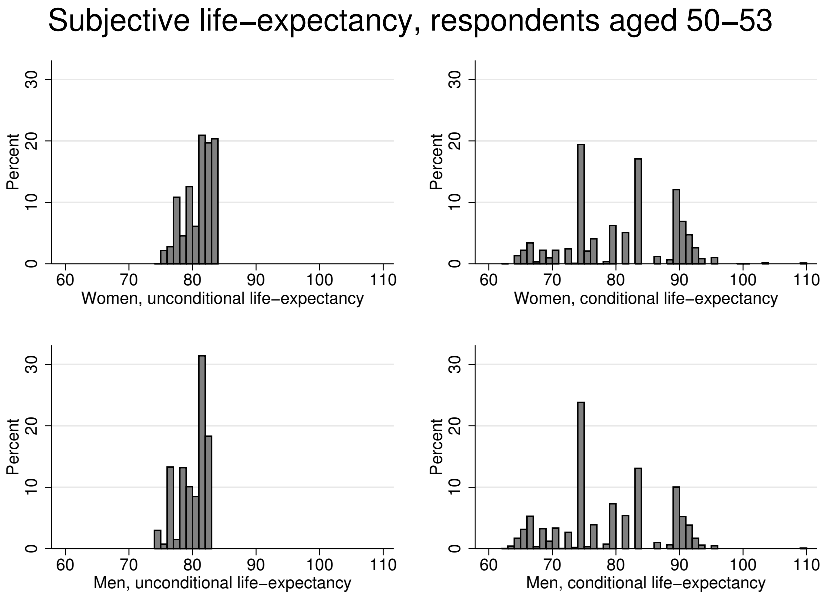


Figure 5.5: Distribution of the conditional life-expectancy (in remaining years) for the respondents aged 50-53, all cohorts combined.

In order to visualize the results obtained at various ages, consider Figure 5.6 presenting the objective and subjective life-expectancy (the number of years a respondent is predicted to survive) for all respondents of our dataset. The figure presents smoothed trends for the median objective survival prediction at each age, and then, smoothed conditional subjective trends for the 10th, 50th, and 90th percentiles, by gender and cohort.

A first striking result is how close the median subjective and median objective lines are for all six sub-figures. As expected from Figure 5.4, median respondents have subjective expectations that are quite close to the objective ones. At first glance, it may therefore be surprising that we find signs of optimism for certain groups of the population. In order to explain this, we have to consider the dispersion of the subjective years. We can see by looking at the 10^{th} and 90^{th} percentiles, for instance, that there seems to be an asymmetry in the answers, and that some respondents are quite optimistic without equivalently pessimistic respondents. The trend seems even more pronounced at later years in life.

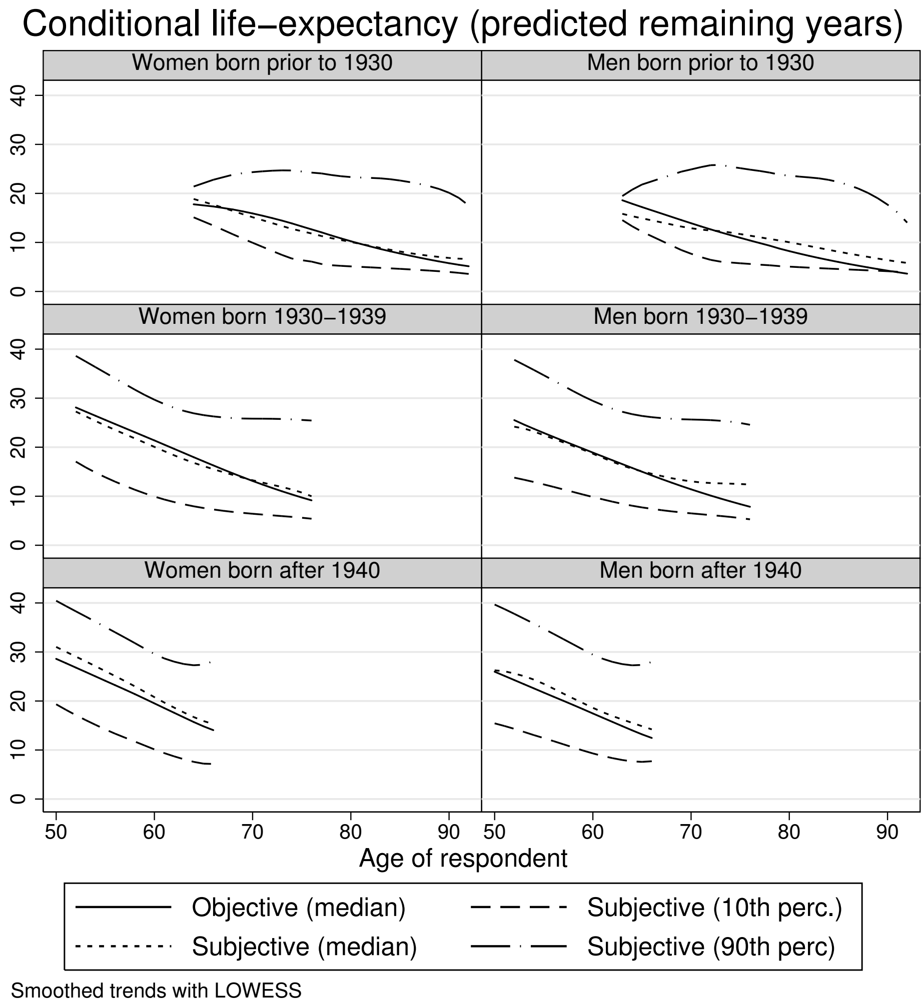


Figure 5.6: Objective and subjective life-expectancy (in remaining years) at various ages, by gender and cohort

5.4.4 *The case of smokers*

We have already mentioned that one of the interesting features of our approach is that it allows us to assess the effect of personal characteristics on expected mortality, even if life-tables are not published specifically for respondents with that characteristic. Let us consider the case of smokers. The accuracy of the expectations of smokers was previously studied by Khwaja et al. (2007) who found, using the HRS and relying on a similar comparison between subjective and objective probability of survival, that smokers tend to be optimistic concerning their own survival probabilities. To relate this to our results, consider Figure 5.7, presenting conditional subjective survival curves for respondents aged 50 to 53 and from the 1940 cohort, but this time presented separately whether they report ever smoking or not. We again add to the figure the objective survival curve of the median respondent.

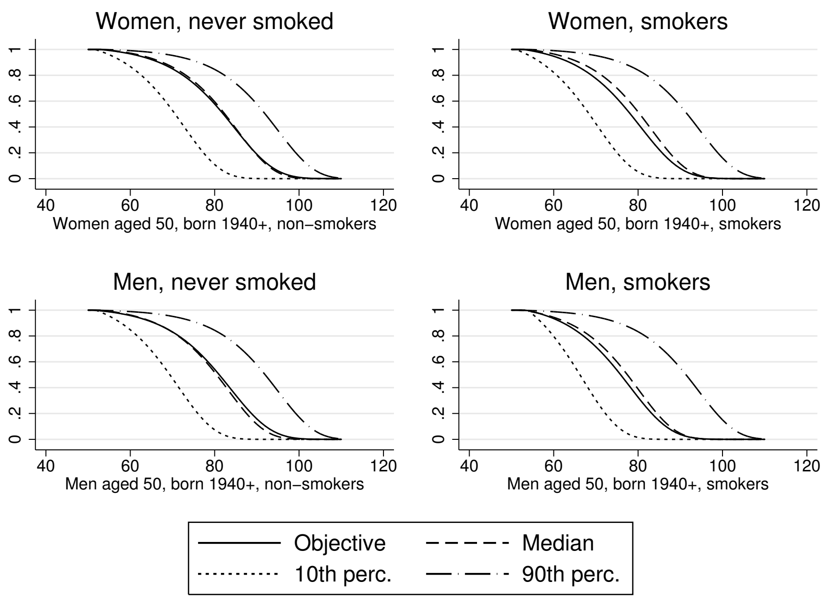


Figure 5.7: Subjective survival curves conditional on the respondents' self-reported probabilities (expected subjective frailty used) presented separately for smokers and non-smokers.

Based on this picture, we would tend to say that median respondents are rational concerning their survival probabilities. This would seem to contradict the finding of Khwaja et al. (2007). However, it does not. To illustrate this, we can can compute the ratio of predicted subjective life-expectancy to predicted objective life-expectancy for every respondent. Let us consider the mean value of this ratio for the four groups of respondents from Figure 5.7. For women who never smoked, we find a value of 1.000 as a mean value for the ratio. This value is 1.024 for men who never smoked. This hints that these two groups are, on average, quite rational, although testing formally the significance of these results is not straightforward, given that we consider the mean of the ratio of two random variables. The picture is different for smokers. The mean value of this ratio is respectively 1.102 and 1.121 for women and men. On average, predicted subjective life expectancy is about 10% higher than predicted objective life expectancy for smokers. Hence, while we find that the median respondent is somehow rational, the distribution is skewed. On this topic, let us mention that Khwaja et al. (2007) compared former and current smokers, and found that those who are no longer smoking tend to be pessimistic, while those still smoking are optimistic. This could explain the result that we obtain.

5.5 CONCLUSION

In this paper, we extend the framework of Gan et al. (2005) so that it allows us to test formally the hypothesis of rational expectations, using subjective survival expectations and observed mortality, and to recover individual subjective survival curves. Our model also explicitly takes rounding behavior into account. Our analysis revealed a general increase of optimism over time, captured by a strong dynamic selection in subjective survival probabilities. We also found that education, gender, and race are important factors in determining optimism, and found evidence that smokers underestimate the effect of smoking. We found a higher mortality among women than is predicted by the life-tables. However, we found that women seem to anticipate this, as median survival predictions are quite rational for that group. Finally, we conclude that the heterogeneity contained in the subjective answers to probability questions is substantial, and unaccounted for by the other characteristics used in our model.

An important problem that must be addressed in future work is the problem of non-response in subjective probability questions. While our approach allows us to control for selection within the sample, this selection problem limits the extent to which we can generalize our results to the whole population. We could extend our current framework to try to predict the subjective expectations of those who refused to answer this question. Additionally, we could exploit the fact that respondents answered the survey multiple times, or even answered more than one question at once, to include a panel dimension in our model.

Finally, and more importantly, we will also consider applications of these individual subjective survival curves to projections of life expectancy and to the understanding of old-age consumption and saving decisions. A good application for our survival curves consists in analyzing the market for annuities or life-insurance. We could assume, for instance, that banks and insurance firms base their pricing on the objective mortality curves, whereas the individuals use their own subjective beliefs concerning their survival probabilities in order to decide whether to annuitize or insure themselves. The subjective optimism that we find would have important implications for these decisions.

5.6 ACKNOWLEDGEMENTS

This project was realized while I was visiting RAND Corporation in Santa Monica, and I should like to thank them for their hospitality. We also want to thank Rob Alessie, Eric Bonsang, Juan Juan Cai, Katherine Carman, Arthur van Soest, and the participants at the 2010 Canadian Economic Association Meeting and the 2011 International Pension Workshop for their insightful comments. Remaining mistakes are ours.

French, E. (2005) The Effects of Health, Wealth, and Wages on Labour Supply And Retirement Behaviour. *Review of Economic Studies*, 72, 395-427.

Gan, L., Hurd, M.D. and McFadden, D. (2005) "Individual Subjective Survival Curves," in *Analyses in Economics of Aging*, ed. by D. A. Wise, vol. I, pp. 377–411. University of Chicago Press, Chicago.

Heitjan, D.F. and Rubin, D.B. (1990) Inference from Coarse Data Via Multiple Imputation with Application to Age Heaping. *Journal of the American Statistical Association*, 85, 304-314.

Heitjan, D.F. and Rubin, D.B. (1991) Ignorability and Coarse Data. *The Annals of Statistics*, 19, 2244-2253.

Hurd, M.D. and McGarry, K. (1993) Evaluation of Subjective Probability Distributions in the HRS. *NBER Working Paper Series*, (4560).

Hurd, M.D. and McGarry, K. (2002) The Predictive Validity of Subjective Probabilities of Survival. *The Economic Journal*, 112, 966-985.

Hurd, M.D. (1989) Mortality Risk and Bequests. *Econometrica*, 57, 779–813.

Kleinjans, K.J. and Van Soest A. (2010) Nonresponse and Focal Point Answers to Subjective Probability Questions. *IZA Discussion Paper*, (5272).

Khwaja, A., Sloan, F., and Chung, S. (2007) The Relationship Between Individual Expectations and Behaviors: Mortality Expectations and Smoking Decisions. *Journal of Risk and Uncertainty*, 35, 179-201.

Lee, R. D. and Carter, L.R. (1992) Modeling and Forecasting U. S. Mortality. *Journal of the American Statistical Association*, 87, 659–671.

Manton, K.G., Gu, X., and Lamb, V.L. (2006) Long-Term Trends in Life Expectancy and Active Life Expectancy in the United States. *Population and Development Review*, 32, 81-105.

Olshansky, S. J., Passaro, D. J., Hershow, R. C., Layden, J., Carnes, B. A, Brody, J., et al. (2005) A Potential Decline in Life Expectancy in the United States in the 21st Century. *New England Journal of Medicine*, 352, 1138-1145.

Vaupel, J.W., Manton, K.G. and Stallard, E. (1979) The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality. *Demography*, 16, 439–454.

Yaari, M.E. (1965) Uncertain Lifetime, Life Insurance, and the Theory of the Consumer. *Review of Economic Studies*, 32, 137-150.

6

USING STATED-PREFERENCE DATA TO ESTIMATE A STRUCTURAL MODEL OF RETIREMENT AND SAVINGS

6.1 INTRODUCTION

It is well documented that populations of developed countries are aging, a trend casting doubt on whether the current pension systems are sustainable. To curb the cost of social security, political reforms are needed around the world. For instance, in the Netherlands, where the dependence ratio is projected to double by 2040 to around 2:5 (see Bovenberg and Gradus, 2008), measures were already taken to reduce the cost of the system by delaying eligibility to social security (known as AOW, in Dutch) by two full years. Predicting how people will react to these reforms is quite challenging. Will they decide to delay retirement by a full two years? Will they save more in order to be able to retire as early as they used to? Or will they, as economic theory would predict in many cases, use a mix of these two solutions in their retirement planning? The life-cycle theory helps us to evaluate behavioral adjustment. To obtain quantitative predictions based on a life cycle model, we need to calibrate or to estimate the preference parameters. Doing so is not a trivial task and usually requires a substantial amount of information on individuals' characteristics, their behavior, and their pension plans.

There is a vast literature on estimation of preferences for retirement based on survey data. A frequently used approach is to assume that a survey respondent evaluates the benefits of retiring at various ages, for instance every year between age 60 and 70, and that the planned retirement age is the one with the highest value. International comparisons of preferences for retirement are available in two volumes edited by Gruber and Wise (1999, 2004) based on the option-value model of Stock and Wise (1990). An alternative literature estimates structural retirement models derived from the life-cycle theory (see for instance Gustman and Steinmeier, 1986; Rust and Phelan, 1997). In many cases, however, the information needed to estimate these complex models, such as the values of assets or detailed descriptions of retirement plans, are not available to most empirical analysts.

Even when these important data are available, one of the main challenges in estimating detailed retirement models is to characterize correctly the choice set of the economic agents. While we may have information on behavior under his actual plan, and while for a given set of parameters we can forecast how he would act under an alternative plan, the challenge for an empirical analyst is to characterize the set of alternative retirement trajectories as perceived by the respondents. For instance, some respondents may not perceive an alternative trajectory as being available to them. In the case of the Netherlands, reaching age 65 is a legitimate reason for dismissal, so respondents may not perceive that retiring beyond that age is within their choice set. Moreover, it is hard to impute a correct expected replacement rate for the feasible ages of retirement that are not chosen by a respondent.

Beyond that, even characterizing the perceived actual retirement trajectory may be quite challenging. There is substantial evidence that many individuals do not understand their own retirement plans (e.g. Gustman and Steinmeier, 2005) or lack the financial knowledge needed to understand complex financial decisions (e.g. Lusardi and Mitchell, 2007). Therefore, when we observe respondents' behavior, we may not be able to know exactly what respondents are reacting to, as disentangling preferences and expectations may turn out to be particularity challenging (see Hurd, 2009, for a review of retirement expectations). Suppose that we observe an individual with very low savings for retirement. It is impossible, based on this observation, to say whether this behavior is due to a strong preference for present consumption, to the individual's expectation that the pension plan will be generous, or to a low preference for leisure. This is particularly problematic if we want to study retirement planning of relatively young respondents who make decisions concerning a very long horizon with substantial uncertainty.

In this paper, I explore an alternative approach to the estimation of retirement models based on the life-cycle theory, namely the use of stated-preference data where respondents are asked to evaluate hypothetical retirement paths. In this case, all the relevant information concerning the hypothetical scenarios is presented to the respondents, so that the choice set is completely observed. The main objective of this paper is to assess how the respondents' stated preferences can be translated in terms of a simple structural model of retirement and savings. The estimated model is then used to simulate a two-year delay in eligibility to social security in line with the proposed policy reform in the Netherlands. Studies based on this idea have been presented before by van Soest et al. (2006) and Voňková and van Soest (2009), but they did not use the stated preference data to estimate a structural life-cycle model. This study partly uses the same data that were used in Voňková and van Soest (2009).

There are many conceptual advantages of the stated-preference approach to estimate the parameters of the proposed life-cycle model. First, I mentioned above that characterizing the full choice set of the respondents is very challenging in a revealed-preference analysis, as there is rarely information available on the suboptimal retirement paths. In a case where the respondents are asked to evaluate multiple retirement paths, I do not have this problem. Second, the fact that empirical analysts can experimentally change the hypothetical scenarios allows us to induce variation that is not naturally present. For instance, it is possible to inquire about multiple replacement rates that do not correspond to the rates a respondent faces. Doing so helps to disentangle the effects of various preferences (e.g. preferences, risk aversion, preference for leisure). Similarly, it would allow us to evaluate which policy reform would be preferred among a list of possible changes without actually implementing these reforms. Third, the hypothetical scenarios allow us to reduce uncertainty (combining the roles of preferences and expectations for the empirical analyst), as the retirement paths presented to the respondents have no uncertainty in the level of retirement income. Finally, the fact that we can ask respondents to evaluate multiple scenarios provides us with richer information on preferences than in the case where only the preferred choice of the respondent is observed.

Economists have usually been reluctant to use subjective data such as stated-preferences. Discussions on this topic are presented, for instance, by Bertrand and Mullainathan (2001) or Dominitz and van Soest (2008). Common objections include the fear that respondents may under- or over-state their preferences for a certain outcome when asked hypothetical questions (mostly in cases were some answers are seen as socially desirable), that they may not be willing to answer the questions honestly, or that they will not think carefully about their answers. These are certainly valid concerns. In the application that I propose here, however, the hypothetical scenarios are easy to understand and I would argue that there is no socially desirable answer. Therefore, I believe that it is worthwhile to investigate the information contained in these stated preferences.

Most importantly, I do not want to argue that stated-preference data should substitute for revealed-preference data. The ultimate aim of this research is to propose a means to complement the information contained in revealed preferences. For instance, the additional data may allow estimation of parameters that are otherwise poorly identified, or we could use stated-preference data to construct additional moment conditions in the estimation of a model of retirement. I therefore believe that it is relevant to investigate the information contained in these answers, keeping in mind that I will want to combine this information with more traditional sources in future research. At this stage, however, my main question is whether this source of information yields plausible estimates of the model if we only rely on it for the estimation.

The remainder of the paper is organized as follows. Section 6.2 describes the data used in this paper. Section 6.3 introduces the economic model, and then explains an econometric approach to estimate the parameters of this model. Section 6.4 presents the results of the estimation. Section 6.5 presents some simulations based on the estimated model, illustrating the implications of the results. Section 6.6 concludes.

6.2 DATA

I first introduce the data I use in this paper, as the information available will dictate the form of the empirical model. The main datasets used are available from CentERdata, an institute affiliated with Tilburg University. The data belong to the CentERpanel, an ongoing internet panel of about 2000 households in a given time period. Respondents to the panel are sampled to be representative of the Dutch population. Measures are even taken to ensure that families without access to the Internet can still be included in the sample.[1]

6.2.1 *Hypothetical scenarios and stated-preferences*

The main dataset used for this research is the *Netspar Pension Barometer Survey*. The *Pension Barometer Survey* is an ongoing high-frequency panel interested, as its name implies, in various perceptions concerning the future of retirement. Among the themes studied via this panel are expectations (e.g. Van der Wiel, 2008, 2009; Bissonnette and van Soest, 2011), satisfaction with pension (e.g. De Bresser and van Soest, 2009), and the perception of the 2008 financial crisis (e.g. Bissonnette and van Soest, 2010). I focus my attention on the yearly survey sent to respondents in the years 2006 to 2008.

In these waves, respondents were asked to evaluate various retirement scenarios using a 10-point scale (1 being "very unattractive" and 10, "very attractive"). Respondents were asked to report what they would think of hypothetical retirement plans if they were offered to them. These scenarios were adapted to the situation of the respondents. For instance, respondents were asked in a prior question to report the number of hours worked weekly, and this information was used in the definition of the number of hours worked until retirement. I analyze three scenarios concerning retirement at ages 62, 65, and 68. For each of these scenarios, respondents were told that they would work the same number of hours as they do now, and were offered randomly a given replacement rate after retirement.[2] After a short introduction explaining the stated preference set-up, the following questions were asked (with everything between square brackets replaced by a relevant value):

1 More information on the CentERpanel is available via the CentERdata website: www.centerdata.nl/en

2 Defined benefits pensions are widespread in the Netherlands, so expressing retirement benefits in terms of a unique replacement rate of $x\%$ is not unusual for our respondents. See Appendix 6.B for a discussion on this topic.

> What do you think about the following possibilities? Please provide answers on a scale from 1 (very unattractive) to 10 (very attractive).
>
> You work [current working hours] till age [*r*], retiring fully at age [*r*] with a disposable pension income of [*x*]% of your last net earnings.

Respondents were assigned randomly a set of replacement rates for these scenarios. For instance, in 2006 or 2007, respondents could receive either low values for all three scenarios (45% for retirement at age 62, 60% at age 65, and 80% at age 68), medium values for these scenarios (50%, 65%, 85%), or high values (55%, 70%, 90%). Values were slightly different in 2008 to allow for additional variation, but respondents would still receive either low, medium, or high values in all scenarios. Over the 3 waves of the panel, respondents therefore faced one of six sets of replacements rates (i.e. either low, medium, or high with the values in 2006-7 or 2008). This experimentally induced variation in replacement rates will play a key role in identifying the preference parameters of the model. Respondents were told to interpret the replacement rate of a given retirement scenario as a substitute for social security and professional pension (the first and second pillar pension). Pension entitlements in the Netherlands are typically presented in this integrated manner, for example in the annual overviews provided by the occupational pension funds (c.f. OECD, 2011).[3]

6.2.2 *Households and their assets*

I allow respondents to accumulate assets to save for retirement. In turn, this implies that I must take into account the accumulated assets at the time respondents evaluate the scenario. While this information is not available in the *Pension Barometer Survey*, it is available for most respondents through the *Dutch National Bank Household Survey* (henceforth DHS), another yearly survey from *CenterData* administered to the same respondents panel. The definition of wealth I use in this paper is simply the sum of the value of different assets included in the DHS. A complete list of the assets included is included in Appendix 6.A.[4] Analysis of these assets can be found, for instance,

3 One of the main advantages of the stated-preference approach used in this paper is the simplicity of the future that respondents have to evaluate. The scenarios do not require institutional knowledge of the pension system or understanding of their current retirement scheme (see for instance Gustman and Steinmeier, 2005). I refer readers interested in comparing the theoretical and hypothetical pensions plans to Appendix 6.B.

4 For this analysis, I excluded housing wealth and mortgage debt, due to two issues. First, there seem to be problems with the measurement of these values in the DHS. Second, including housing would require modifications to the theoretical model

in Alessie, Hochguertel, and van Soest (2002), Euwals, Eymann, and Börsch-Supan (2004) or Bissonnette and van Soest (2010).

One of the limitations of the paper is that the scenarios were phrased in terms of individual behavior rather than in terms of household behavior, while the value of assets is defined at the household level. Ideally, scenarios would also explicitly describe the spouse's behavior. Hence, due to data limitations, I cannot estimate a full model of retirement for the couples (e.g. Gustman and Steinmeier, 2004).

Moreover, information in the *Pension Barometer Survey* and in the *DHS* is not always available for each member of the household. In order to include as many respondents as possible, I decided to divide equally household wealth and income among the two household members.[5] This will probably underestimate the retirement cost of the high earner and overestimate the retirement cost of the low earner. However, not accounting for other household members' income would clearly overestimate the consumption of the highest earner and underestimate consumption of the lower earner. A better analysis of household behavior is a step that will require further investigation and data, in future research. In the current study, I restrict the sample to singles and couples where both members are observed. Then, each respondent is allocated the mean income within the couple and half of the total values of the assets.

After restricting the respondents to singles and "half-couples" aged less than 61, merging the dataset with the DHS, and removing observations with insufficient information, I retain a sample of 607 respondents, forming a panel of 963 respondent-year observations.

6.2.3 *A descriptive analysis of stated-preferences*

In this subsection I present a descriptive analysis of the stated-preference data for the selected households. Although respondents were asked to evaluate the scenarios on a 10-point scale, the information that I exploit in this paper is the preference relation between the three scenarios. Hence, if a respondent evaluated the scenario with retirement at 62 with a 4, the one with retirement at 65 with a 7, and the one with retirement at 68 with a value of 6, my focus will be that the value of the second scenario, denoted V_{65}, is larger than the value of the scenario at 68, V_{68}, in turn larger than the value of retirement at 62,

going beyond the scope of this paper. If houses are considered as assets, then I would either need to establish a utility benefit of holding a house (or a housing cost for the respondents who do not own a house) or assume that houses are assets with a very high rate of return to compensate for mortgage cost. The implicit assumption in my model is that agents may derive utility from owning a house and that the price of this good is the mortgage paid over the years, part of the Hicksian good.

5 This assumes that all consumption is private; an alternative would be to assume that all consumption is public, leading to twice as high consumption amounts for both partners. This factor two does not change anything in the empirical results since transitions into widowhood etc. are not considered.

V_{62}. As I explain in detail in Section 6.3.2 below, I do so in order to avoid modelling the reporting procedure respondents used to map their value to a ten point-scale. This procedure has the advantage of being robust to systematic pessimism or optimism in reporting.

Table 6.1 presents the distribution of the rankings as observed in the sample. As explained above, in 2006 and 2007, respondents were allocated randomly one of three sets of replacement rates and the values of these replacement rates were changed in 2008. Therefore, over the years, respondents had to evaluate scenarios with one of six sets of replacement rates. Each of the first six columns of the table corresponds to the six different sets of replacement rates that were presented to the respondent. The seventh column pools all respondents, irrespective of their replacement rates. This table ignores the fact that some respondents are observed more than once over time.

Let us consider the last column first. We see that across the scenarios, the ranking observed most frequently is such that retiring at 68 is preferred to retiring at 65, in turn preferred to retiring at 62. If we focus on this result only, it would indicate that late retirement is the favored option. It is not exactly so. Let us consider only cases where the preferred option is known (respondents from the first nine rows) (excluding the 139 respondents for whom only the least preferred option is known and the 135 respondents for whom no ranking is available). Out of these 689 respondents, 21.9% prefer the scenario with retirement at 62, 39.3% prefer retirement at 65, and 38.8% prefer late retirement. Hence, if we exclude the ties among the favoured options, about as many respondents prefer standard and late retirement. Among the cases with a tie between the two preferred scenarios, 59.0% expressed that early retirement was the least preferred option, 8.6% chose standard retirement and 32.3% chose late retirement. While it is clear that early retirement is the least preferred option among the three scenarios, it seems that the fraction of respondents preferring standard and late retirement are about the same.

As expected, variations in the generosity of hypothetical pension plans seems to have an effect on the stated choices. We also see that in the first column, where early replacement rates are the lowest, respondents express a preference for late retirement, chosen by 48% of the 163 respondents for whom the preferred option is known. To illustrate the impact of a variation in replacement rates, consider the sixth column with more generous replacement rates for all scenarios. In this case, most respondents prefer standard retirement (51% of 164 respondents for whom the preferred option is known).

Table 6.1: Distribution for the ranking of retirement scenarios by set of hypothetical replacement rates.

	Replacement rates when retiring at 62/65/68						
	45/60/80	50/65/75	50/65/85	50/70/85	50/75/95	55/70/90	Pooled
$V_{62} > V_{65} > V_{68}$	14	14	16	9	5	17	75
$V_{62} > V_{68} > V_{65}$	3	0	1	0	0	2	6
$V_{65} > V_{62} > V_{68}$	13	7	13	16	12	32	93
$V_{65} > V_{68} > V_{62}$	16	3	11	13	13	21	77
$V_{68} > V_{62} > V_{65}$	4	0	7	1	0	3	15
$V_{68} > V_{65} > V_{62}$	56	10	44	11	14	43	178
Total - Full ranking	106	34	92	50	44	118	444
$V_{62} > V_{65}, V_{62} > V_{68}$	18	5	28	7	4	8	70
$V_{65} > V_{62}, V_{65} > V_{68}$	17	9	27	9	8	31	101
$V_{68} > V_{62}, V_{68} > V_{65}$	22	7	30	5	3	7	74
$V_{62} < V_{65}, V_{62} < V_{68}$	15	9	12	4	15	27	82
$V_{65} < V_{62}, V_{65} < V_{68}$	3	3	3	1	0	2	12
$V_{68} < V_{62}, V_{68} < V_{65}$	16	5	10	2	1	11	45
Total - Partial ranking	91	38	110	28	31	86	384
Info on preferences avail.	197	72	202	78	75	204	828
No ranking available	37	12	40	6	11	29	135
Total	234	84	242	84	86	233	963

6.2.4 *Additional data to construct the future*

While a lot of information is available through the two sources described above, solving the life cycle optimization problem requires some additional assumptions, in order to construct the full income paths and to describe their mortality risks. I describe these elements here. Please note that all variables are described in terms of real values, as I do not consider inflation in the model.

Given that the pension income is described in terms of the replacement rate of the yearly wage, the major challenge is to construct sensible wage paths for the respondents. Ideally one would probably ask the respondents about their own expected last earnings – information that will be available in subsequent waves of the *Pension Barometer Survey* but is not yet available in the waves used here. An alternative is to rely on an external prediction of income growth. I use this approach, and calibrate my model based on the results of Knoef, Alessie, and Kalwij (2009), who present microsimulations of income growth in the Netherlands. I use their finding that they expect the growth of real income of the 50-64 to be 0.6% between 2008 and 2020. I therefore impute this income growth to every individual. A more refined projection by characteristics such as education or by quantile of income is left for future research.

Another problem concerns the expected real interest rate. For this part, I follow standard practice and present a sensitivity analysis using various assumed values. I hence assume that respondents all expect the same interest rate.

Finally, the theoretical model includes the survival probabilities to various ages. The best approach here would again be to consider heterogeneous subjective survival ages, using information revealed by the respondents (e.g. Gan, Hurd, and McFadden, 2005). For the sake of simplicity, however, I consider that all respondents of a given age and gender have similar mortality expectations. To determine the probability of survival, I consider the cohort life-tables as forecasted by *Statistics Netherlands* (in Dutch: CBS).[6]

6.3 LIFE-CYCLE MODEL

The main objective of this paper is to estimate a structural model of retirement and savings based on the valuation of retirement scenarios

6 The life-tables contain forecasts of mortality until 2060, meaning that projections of survival probabilities are not known toward the end of the life course for respondents born later than 1960. In order to circumvent this problem, I assign the last predicted survival probability to younger cohorts. For instance, the last predicted probability of survival to age 85 is for a respondents born in 1975. Respondents born after that year are assigned the same probability of survival as those born in 1975. Given that these later periods are quite substantially discounted in the model, this approach should have only a limited impact on the estimation results.

described in the previous section. I present in this section the approach I use to do so. I first introduce the underlying economic model, then, discuss the econometric approach used to estimate the parameters of the theoretical model.

6.3.1 *The economic model*

The problem faced by the agent in the model is fairly common in the life-cycle theory (see, e.g., Browning and Lusardi, 1996; Browning and Crossley, 2001). Let us denote the consumption, leisure, and accumulated assets at the end of period t by C_t, L_t, and A_t respectively. In this paper, L_t can only take the values 0 or 1, as respondents either work or are retired. Extension to partial retirement would be straightforward, as discussed by Gustman and Steinmeier (2005). Additionally, let us denote s_{at} the probability to survive from age a to age t. We consider an intertemporally separable utility function given by

$$U_a = \sum_{t=a}^{T} s_{at} \rho^{t-a} u(C_t, L_t) \tag{6.1}$$

where ρ is a discount factor to capture preference for present consumption. For the sake of simplicity, I assume that respondents have a probability of survival to age 68 equal to 1. After that, there is a positive probability of dying at each time period. This assumption is made to avoid additional assumptions concerning the credit market under an uncertain lifespan (see Yaari, 1965, for an early example). In a case where agents could borrow money with uncertain survival status, the interest rate would need to be a function of the survival probabilities of the agents, meaning that the interest paid on a one-year loan would increase as a respondent ages.

Let ω_t be the wage for time period t, $\pi_t(r)$ the pension amount received at time t for a given retirement age r, and let ι denote the real interest rate. Respondents face the following constraints:

$$(1+\iota)A_{t-1} + \omega_t(1-L_t) + \pi_t(r)L_t \geq A_t + C_t \qquad t = a...T \tag{6.2}$$

$$A_t \geq 0 \qquad t = r...T. \tag{6.3}$$

The first constraint is a conventional budget constraint. The second constraint states that while respondents are allowed from borrowing against their future wages, I do not allow them to borrow against their future pension income. It is often assumed that respondents must hold positive asset values at each time period. However, I observe many respondents with debts in the current time period. I therefore decided to allow respondents to borrow against their future wage. As mentioned above, however, I decided to prevent respondents to borrow in the years with uncertain survival status in order to keep the credit market simple. I therefore make the simplifying assumption that respondents enter retirement without debts.[7]

One of the implicit assumptions of this model is that there is no bequest motive (see, e.g., Hurd, 1989), and that hence:

$$A_T = 0.$$

There may still be some accidental bequests should a respondent pass away while holding wealth. I assume, however, that a respondent derives no utility from this bequest.

As for the utility function, I use a separable utility function with constant relative risk aversion for consumption. This function was chosen for the sake of comparability with structural estimations reviewed above. Generally speaking, it has the form:

$$u(C_t, L_t) = \frac{1}{\gamma} C_t^{\gamma} + \lambda L_t \tag{6.4}$$

where the parameters γ and λ vary across individuals according to respondents' characteristics, as discussed below in Section 6.3.2. When specified in this way, the coefficient of relative risk aversion for consumption is equal to $1 - \gamma$. With this specification, the elasticity of intertemporal substitution is given by $\frac{1}{1-\gamma}$. For $\gamma = 0$ the model becomes a model with logarithmic utility.

6.3.1.1 *Solving the agent's problem*

In order to determine the ranking of the scenarios for a given respondent, we need to determine the optimal asset allocation over time for a given set of parameters and for given initial financial wealth, retirement age, and incomes before and after retirement. In the simple model considered here, the only source of uncertainty is the survival status of a respondent in the future. The problem is therefore computationally easy to solve and is akin to a model where the future is discounted with a time varying factor.

Finding the optimal solution for this simple problem can be achieved by solving the system of first order conditions. In a case where there is no borrowing constraint, the problem consists in solving a simple system of linear equations. Doing so allows us to find an analytical solution to the consumption at the initial time period that can, in turn, be used to solve for the other time periods. However, I want to prevent the agents from borrowing against their pensions. To deal with this condition, I first solve the problem for all time periods from a to T without taking into account the constraint that financial assets must

7 To some extent, this assumption seems to be in line with what is observed in the dataset, in which I observe 12% of working individuals and 5% of retired individuals with negative assets value.

have a non-negative value after retirement. Should the solution violate the non-negativity condition, I set A_{T-1} equal to 0 and re-solve the system for the first $T - a - 2$ decisions. I repeat this procedure until I find the first admissible solution.

The assumptions made in the current paper lead to a very simple solution to the optimization problem faced by the agents. In future research, adding more sources of uncertainty, such as income or health uncertainty, would clearly enhance the interest of the model. Indeed, the stated-preference approach presented here would be an interesting avenue to study these sources of uncertainty. We could, for instance, vary experimentally the risk of unemployment or of deteriorating health, and measure how respondents react.

6.3.2 *The econometric model*

While working with revealed preferences often forces econometricians to estimate a model based only on the alternative providing the highest utility to a respondent, the use of stated preferences allows us to use more information concerning preferences. In the analysis of the same dataset, Voňková and van Soest (2009) exploited the fact that respondents used a response scale from 1 to 10 as a measure of the level of preferences, and estimated a multivariate ordered probit. I use an alternative approach here and only consider the ordering of the answers, ranking alternatives according to preference relations. My main motivation for doing so is that there is no straightforward way to map the value of the utility function to a valuation system from 1 to 10. Consider for instance the case of two identical respondents, but at different ages. Even in the absence of a stochastic element in the model, these respondents would have different values for the various scenarios, although they are likely to rank the scenarios in the same way.

The implicit assumption made is that reported answers are an unspecified monotonic transformation of the intrinsic value held by respondents for a particular scenario. In some cases, the transformation leads respondents to report the same value for various scenarios. In such cases, the answers are considered as uninformative. I discuss below how I deal with such observations.

I assume a simple model of random utility where the value of a scenario at a given age a, with retirement age r, is given by:

$$V_r(A_{a-1} \mid \omega_a, \boldsymbol{x}, \alpha) = \max_{A_a} \Bigg(\frac{1}{\gamma(male)} C_a^{\gamma(male)} + \exp(\boldsymbol{x}'\boldsymbol{\beta} + \alpha) L_a \\ + s_{a(a+1)} \rho \, \mathrm{E}\big(V_r(A_a)\big) \Bigg) + \sigma_\varepsilon \varepsilon_r. \qquad (6.5)$$

Note that when considering a given retirement scenario r, the value of L_t is equal to 1 of $r \geq t$ and to 0 otherwise.

Observable characteristics enter the model through the parameters γ and λ in equation (6.4). The parameter of risk aversion γ varies only with the gender of the respondent. This parameter has a value of γ_0 for women and $\gamma_0 + \gamma_{\text{male}}$ for men. On the other hand, the marginal utility of leisure $\lambda = \exp(\boldsymbol{x}'\boldsymbol{\beta} + \alpha)$ can vary with observed characteristics in a vector $\boldsymbol{x}$, including variables relating to education, gender, age, and marital status. Age is included by means of a quadratic function using the age of the respondents in 2006 minus 22 (the age of the youngest respondent in 2006). Note that age is age at the time of the survey – age in this model is a time-invariant regressor aimed at capturing cohort effects rather than age effects. Marital status is included by means of a dummy variable with value 1 if the respondent belongs to a couple and 0 otherwise, and an additional term interacting this marital status dummy with the dummy *male*. Table 6.2 presents basic descriptive statistics for these independent variables.

Table 6.2: Descriptive statistics of independent variables at the first time period a respondent is observed (N=498)

	Mean	Median	Std. Dev.
Age	42.24	42.00	9.45
Man	0.50	0.00	0.50
%	0.00	1.00	
Educ. Med.	0.33	0.00	0.47
Educ. High.	0.53	1.00	0.50
Partner	0.51	1.00	0.50
Net yearly wage	19,432.18	18,996.00	5,821.72
Net assets/debts	36,710.16	14,167.54	72,714.52

We see that the resulting sample has about as many men as it has women, and as many singles as non-singles. As we can see from the education dummies, the sample is highly educated, with most respondents having what is considered as a high level of education. This seemingly high level of education may partially be explained by the fact that we focus on employed respondents, but nevertheless hints that selection may be a problem.

The model includes two types of unobserved terms: the unobserved heterogeneity α in preference for leisure and the retirement specific error term ε_r. The error terms ε_r capture mood effects and reporting error, and represent variations in preferences at the moment the respondent is asked to evaluate the scenarios. For the sake of simplicity and tractability, I assume that the error terms ε_r are i.i.d. distributed with an extreme value type 1 distribution, leading to an econometric model of the logit family. This error term does not have a persistent effect for the expected value function at time $a+1$ and has an expected value of 0 in the future. Persistence in preferences is captured only by the means of heterogeneity in preference for leisure. I include respondent specific random-effects by means of the term α. I exploit the panel dimension of the data, and assume that this term α is constant over the waves that a respondent answered the survey. I assume that this term is normally distributed with standard error σ_α, to be estimated. It follows that a systematic preference for early retirement over the waves of the survey is captured by means of a high value of α, while reporting once a high preference for this scenario would be captured by a high draw of the error term at this period. Finally, in the estimation procedure, I constrain the parameter ρ to be between 0 and 1, as I enforce the constraint that respondents have a preference for present consumption.

Note that the presence of random-effects α and of the error terms ε_r does not affect the way I solve the optimization problem faced by the agents, presented above. Given that I use a separable utility function, the optimal consumption for a given retirement scenario is not affected by the value of λ, implying that the heterogeneity term α or any of the covariates included in $\boldsymbol{x}$ (with the exception of gender) will not affect optimal asset allocation. Similarly, given that ε_r is interpreted as a non-persistent reporting error term, the value of this parameter does not affect the optimal solution to the maximization problem.

6.3.3 *Likelihood*

As already mentioned, the estimation of the model relies on the ranking of the scenarios. This approach contrasts with the usual estimation of life-cycle models of this type that tend to rely either on variation in consumption over time or on moment conditions based on observed retirement. Given the assumptions made above, the model can conceptually be characterized as a rank-logit model (see Beggs, Cardell, and Hausman, 1981) with random-effects in the leisure parameter. In order to allow for these random-effects, the model is estimated by maximum simulated likelihood. In what follows, I will consider the term of unobserved heterogeneity as given to keep the presentation simpler. Moreover, for the sake of exposition, I define:

$$\begin{aligned} V_r(a) &= V_r(A_a \mid \omega_a, \boldsymbol{x}, \alpha) \\ V_r^*(a) &= \frac{V_r(a)}{\sigma_\varepsilon} - \varepsilon_r. \end{aligned}$$

Suppose we have self-reported values, on a 10-point scale, of three retirement paths denoted 1, 2, and 3. In the best case, it is possible to determine the full ranking of the three scenarios. Say that a respondent scores retirement path 1 better than path 2 and path 2 better than path 3. In such a case, the likelihood contribution of is given by

$$L(V_1(a) > V_2(a) > V_3(a)) = \frac{\exp V_1^*(a)}{\sum_{z\in\{1,2,3\}} \exp V_z^*(a)} \frac{\exp V_2^*(a)}{\sum_{z\in\{2,3\}} \exp V_z^*(a)}. \tag{6.6}$$

It may happen, however, that a respondent gives the same value to two or three scenarios, and hence does not reveal the full ranking. In such cases, an additional assumption is needed in order to use the rank-logit model, namely that the ties are exogenous, and do not depend on respondent characteristics, such as ranking capabilities (see Fok, Paap, and van Dijk, 2010, for an example with unobserved heterogeneity in capabilities). The approach I follow in cases of partial ranking is to either maximize the probability that a scenario would be preferred to the other two or the probability that a scenario would be ranked last according to the information available.

Suppose that the respondent prefers 1 to 2 and to 3, but did not express a preference between 2 and 3, then, the contribution is given by:

$$L(V_1(a) > V_2(a), V_1(a) > V_3(a)) = \frac{\exp V_1^*(a)}{\sum_{z\in\{1,2,3\}} \exp V_z^*(a)} \tag{6.7}$$

which is simply be the likelihood of a multinomial logit. Then, in a case where 1 and 2 are preferred to 3, but where there cannot be discrimination between 1 and 2, the contribution is given by:

$$\begin{aligned} L(V_1(a) > V_3(a), V_2(a) > V_3(a)) = \\ \frac{\exp V_1^*(a)}{\sum_{z\in\{1,2,3\}} \exp V_z^*(a)} \frac{\exp V_2^*(a)}{\sum_{z\in\{2,3\}} \exp V_z^*(a)} \\ + \frac{\exp V_2^*(a)}{\sum_{z\in\{1,2,3\}} \exp V_z^*(a)} \frac{\exp V_1^*(a)}{\sum_{z\in\{1,3\}} \exp V_z^*(a)}. \end{aligned} \tag{6.8}$$

This expression corresponds to the sum of the two combinations associated with the scenarios in which 3 is ranked last.

Obviously, a respondent who gives the same value to all scenarios would contribute a constant term to the likelihood and is omitted for the maximization.

6.4 ESTIMATION RESULTS

Estimation results are presented in Table 6.3. Results are presented for two groups: singles only in the left half of the table, and singles and "half-couples" (as described in Section 6.2.2) in the right half. Then, for each of these two groups, the model was estimated using a 1%, 2% and 3% interest rate.

Table 6.3: Estimation results

	Singles only			Singles and "half-couples"		
	1% interest	2% interest	3% interest	1% interest	2% interest	3% interest
ρ	0.965***	0.956***	0.949***	0.968***	0.964***	0.960***
	(0.013)	(0.022)	(0.011)	(0.007)	(0.006)	(0.006)
γ_0	-0.197***	-0.160***	-0.179***	-0.160***	-0.140***	-0.145***
	(0.050)	(0.056)	(0.060)	(0.043)	(0.046)	(0.049)
γ_{male}	0.023	0.015	0.009	0.018	0.013	0.010
	(0.026)	(0.026)	(0.026)	(0.025)	(0.015)	(0.015)
β_{constant}	-1.350***	-1.175**	-1.570***	-1.046**	-0.980**	-1.217**
	(0.145)	(0.552)	(0.589)	(0.430)	(0.455)	(0.485)
β_{male}	0.145	0.076	0.020	0.074	0.032	0.001
	0.280	(0.281)	(0.281)	(0.170)	(0.171)	(0.173)
$\beta_{\text{educ. med.}}$	-0.030	-0.034	-0.044	-0.013	-0.012	-0.014
	0.067	(0.069)	(0.073)	(0.047)	(0.049)	(0.051)
$\beta_{\text{educ. high.}}$	-0.148**	-0.152**	-0.174**	-0.110**	-0.110**	-0.118**
	(0.063)	(0.066)	(0.070)	(0.044)	(0.046)	(0.049)
β_{partner}				0.092*	0.091*	0.096*
				(0.051)	(0.053)	(0.056)
$\beta_{\text{partner}\times\text{male}}$				-0.020	-0.022	-0.025
				(0.067)	(0.070)	(0.074)
$\beta_{\text{age}/10}$	0.110	0.192	0.293	0.157*	0.210**	0.306***
	(0.152)	(0.163)	(0.184)	(0.093)	(0.099)	(0.113)
$\beta_{\text{age-sqr}/1000}$	-0.366	-0.516	-0.671	-0.451**	-0.540**	-0.703***
	(0.360)	(0.377)	(0.414)	(0.219)	(0.233)	(0.258)
σ_λ	0.211	0.209	0.214	0.232	0.237	0.249
σ_ε	0.095	0.107	0.075	0.129	0.139	0.121
Log-lik.	-573.406	-575.552	-576.854	1148.562	1152.500	1156.237
N. Ind.		242			498	
N		411			828	

Standard errors between parentheses

Variable *age* is age in 2006 minus 22

The estimate of the parameter ρ is constrained to be between 0 and 1

The first interesting result is that the discount rate obtained under the various specifications was estimated precisely. The estimated values of the discount parameter did not change much, staying between 0.95 and 0.97 in all specifications. These values seem plausible and are well in line with calibrated values in the studies referred to above. Moreover, in all cases, the estimated values of the parameter γ are also plausible and are in line with findings in the literature For instance, most of the estimates obtained for the larger sample have values around -0.15, leading to a coefficient of risk aversion of 1.15, which would be in the lower spectrum of estimation from other studies (see Blau and Gilleskie, 2006, for some examples). As a comparison, Hurd (1989) estimated a coefficient of relative risk aversion equal to 1.12, very close to the values obtained with the larger sample, and Rust and Phelan (1997) obtained a value of 1.072.

The estimations did not reveal much about the determinants of preference for leisure. The results including only singles, presented in the first three columns, are generally quite imprecise. The only personal characteristics with a significant effect is the dummy capturing the effect of higher education. Adding respondents with a partner leads to slightly different conclusions. First, note that the dummy for whether a respondent is in a couple or not is significant at 10%-significance level (with all p-values smaller than 6%). I now find a significant age trend. If we consider, for instance, the specification with 2% interest rate, the quadratic trend reaches its maximum value at 39 years old, and decreases afterward. For most of the respondents, this effect is of economically small importance. The effect becomes more important for older respondents, which may be partially explained by a selection issue. Remember that the analysis includes only working respondents. Given that early retirement being important in the Netherlands, we would expect that the older respondents are those with lower value for leisure, everything else being held constant.

6.5 simulations

Due to the structural nature of the model, I can use the results to predict the behavior of respondents under various scenarios. In this section, I therefore explore the impact of various hypotheses on actual behavior, paying particular attention to the possible reaction to a delay of 2 years of eligibility to social security.

For the sake of comparison with previous studies, I present here an analysis with a 2% real-interest rate. Parameters estimated with this interest rate and with the large sample including couples were used for the simulation. Given the current economic situation at the time of writing these lines, the real interest rate is probably larger than it should be.

6.5.1 *Initial results*

Consider the original scenarios proposed to the respondents. For each scenario, I can predict the probability that a respondent would prefer this scenario over the others. To do so, random draws of the term of unobserved heterogeneity α are taken, and the mean probabilities to select each scenario conditional on these draws are computed.

Consider Simulation (1) in Table 6.4, presenting the average probability to retire at a given age given the heterogenous replacement rates presented to the respondents. We see that the model predicts large heterogeneity in retirement behavior. While the median and modal retirement age would be 65, the proportion of respondents that would choose to retire at 62 or 68 is substantial. Based on the discussion presented in Section 6.2.2, it is not surprising that the model predicts about the same proportion of respondents with preference for retirement at 65 and at 68. However, the share of preference for early retirement seems quite high, although not entirely off the mark. The results presented here are in line, at least qualitatively, with the description of the Dutch pension system made by Capretta (2007). However, these results are at odds with Voňková and van Soest (2009), who find a strong preference for standard retirement at 65. The difference in results between this paper and theirs could be due to a different subsample, focus on different stated-preference questions or to the method itself.[8]

In these scenarios, each respondent faced different replacement rates, meaning that respondents' preferences were not directly comparable. In what follows, I will consider alternative simulations where respondents are offered comparable replacement rates.

6.5.2 *Reference scenario and actuarial neutrality*

For the remainder of the section, I want to offer comparable replacement rates to all respondents conditional on their retirement at a given age. I start from a common scenario where respondents are offered a 70% replacement should they retire at 65. This scenario will serve as a benchmark. I use this information to compute replacement rates under alternative scenarios. One important concept in the determination of

8 A quick analysis of the data hints that the variation in the subsample is indeed an important factor. For instance, an analysis of the distribution of answers on the 10-point scale reveals that the respondents included in my sample do not exhibit the strong preference for standard retirement over early retirement that is shown in the histograms of Figure 1(c) in Voňková and van Soest (2009). However, this does not seem to be enough to justify the difference, and further investigation using their estimation and simulations method would be needed to disentangle possible sources of difference.

Table 6.4: Simulation results: Initial results, actuarial neutrality and delayed AOW

	(1)		(2)		(3)		(4)		(5)	
	Hyp. scenarios.		Act. neutral (author's)		Act. neutral (author's)		Act. neutral (V & vS, 2009)		Delayed AOW	
Age	Rep. rate	%	Rep. rate[a]	%	Rep. rate[a]	%	Rep. rate	%	Rep. rate[a]	%
61					52.47	18.83	53.45	21.74	48.79	16.66
62	45/50/55	27.35	56.31	43.79	56.31	13.07	56.95	13.75	52.48	11.78
63					60.48	10.78	60.80	10.57	56.47	9.96
64					65.03	9.74	65.14	9.28	60.83	9.27
65	60/65/70/75	36.23	70.00	29.68	70.00	9.19	70.00	8.72	65.59	9.07
66					75.45	8.80	75.00	7.86	70.80	9.06
67					81.45	8.38	81.16	7.81	76.55	9.02
68	80/85/90/95	36.40	88.08	26.53	88.08	7.83	87.71	7.32	82.91	8.85
69					95.46	7.11	95.17	6.79	89.99	8.46
70					103.44	6.27	103.53	6.15	97.82	7.86
Mean age ret.	65.3		64.5		64.6		64.4		64.9	

[a] Mean replacement rate over the sample

the replacement rate is the annuity factor for a given retirement age r, given by:

$$AF(r) = \sum_{t=r}^{100} \frac{s_{rt}}{(1+\iota)^{t-r}}. \tag{6.9}$$

Remember that both income at 64 and survival probabilities will vary by respondents, however this is omitted from the notation for ease of presentation.

I assume that pensions are defined in terms of the last wage, as was the case in the hypothetical scenarios. Following a description in the OECD guide for pensions, the replacement rate at 65 is defined as 1.75% times the number of years worked, assuming a full career of 40 years. For each respondent, I compute the pension wealth at age 65 for a given retirement age given by:

$$PW_{65}(r) = 1.75\% \, (r-25) \, AF(65) \, \omega_{64} \tag{6.10}$$

which obviously implies that the value of the pension wealth at 65 for someone retiring at 65 is given by:

$$PW_{65}(65) = 70\% \, AF(65) \, \omega_{64}. \tag{6.11}$$

The scheme used here has a strange contribution pattern, in the sense that the contribution's value at age 65 is the same independently of the age at which it was made, and that hence, the real value contributed at each time period is increasing over time.

I first want to compute various scenarios under actuarial neutrality, taking into account both the increase in pension wealth due to the contributions over time by and mortality risk. In the current setting, the actuarially fair replacement rate is given by:

$$reprate_r = \frac{(1+\iota)^{r-65} \, PW_{65}(r)}{s_{65,r} \, AF(r) \, \omega_{r-1}}. \tag{6.12}$$

Note that $s_{65,r}$ is equal to one if $r < 65$ in this model where there is no mortality between 62 and 68.

Column (2) in Table 6.4 presents a simulation of retirement scenarios with the same retirement ages that were used before, but with actuarially neutral replacement rates. To simplify the presentation, the replacement rates presented in the table are the average of the replacement rates offered to each respondent for a given retirement age. On average, these rates are close to the most generous hypothetical scenarios presented to the respondents. These rates would lead to a substantial fraction of the workers retiring at age 62, as almost half of them would be predicted to retire at that age under this progression.

Another way to look at this is that the rates that were used in the scenarios were encouraging work considerably. This could actually be the case. Using a 70% replacement rate at age 65 and actuarially neutral rates for ages 62 and 68 is comparable to one of the most generous sets of replacement rates used in the scenarios, as seen in the 6th column of Table 6.1. In the simulations, agents react to this by retiring earlier. Compared to the values observed in Table 6.1, however, it seems that the model predicts a higher share of respondents preferring early retirement than we observe in the data.

To allow for more flexibility in retirement, Simulation (3) allows a respondent to retire at every age between 61 and 70. The average retirement age in this model is predicted to be 64.6 years.

While the fact that each respondent receives a different set of replacement rates enforces the concept of actuarial neutrality, this approach to pension plans is not realistic. An alternative approach would be to use the same replacement rates for all individuals. In order to do so, I realized two simulations. First, I supposed that each respondents would receive the average replacement rate, as seen in Simulation (3) of Table 6.4. Results are very similar to the one already presented, leading to a decrease in the retirement age of 0.03 years (12 days), and are not reported in the Table. I was also interested in using the OECD replacement rates used by Voňková and van Soest (2009), in order to compare my results to theirs. This was done in Simulation (4), allowing us to assess the impact of variations in replacement rates. We see that using this rate leads to an average retirement age of 64.4, a diminution of 0.2 years, about two months, compared to Simulation (3). This would indicate that the agents in the model are very sensitive to seemingly small variations in replacement rates.

6.5.3 *Delayed eligibility to public pension*

The Dutch government, as most governments in developed countries, is in the process of reforming social security in order to curb future costs. Among the policies evaluated, a progressive increase of the eligibility age to AOW from age 65 to 67 is the most likely element of reform. I mentioned above that public and private pensions are often integrated in the Netherlands, with the aim of offering a given gross replacement rate to a worker retiring at 65. Hence, a shock to social security like this proposed pension reform is not straightforward to implement in my simulations, as I must make additional assumptions concerning how the occupational pensions schemes would adjust.

The approach I consider is to interpret the pension reform as a shock to the value of the pension wealth at 65 that would be absorbed by the occupational pension. I simply remove two years of social security from the pension wealth computed in Equation 6.11 and compute new replacement rates based on Equation 6.12. Given that the scenarios are expressed in terms of net income, I approximate the value of social security to be the value received by singles in 2008, and assume that everyone pays on this amount the lowest rate of income tax. This leads to a gross yearly value of 12,718 euros on which a 33.60% tax is paid. Considering the delay of two years proposed, the net total reduction of $PW_{65}(65)$ is therefore 16,724 euros.

This last simulation, numbered (5), is presented in Table 6.4. We see, for instance, that the average age of retirement increases by 0.3 years, an increase of four months. The impact of this delayed eligibility to social security may look quite small at first glance. It has to be stressed that the Netherlands is a country where individuals accumulate a large pension wealth during their lifetime. According to the most recent OECD report on pensions available at the time of writing (cf OECD, 2011), the median Dutch man accumulates 12.8 times his net income in pension wealth while the median Dutch woman accumulates 14.6 times her net income. These numbers are among the highest in the countries reviewed in the publication. Thus, a reduction by two years of public pension on the total pension wealth does not have an effect as strong as one might expect.

An important aspect that must not be overlooked is that the reform is expected to have different effects on different individuals. For respondents with lower income, for instance, social security represents a larger share of the pension wealth than for richer individuals. Other factors may affect behavior. Younger respondents have more time to adjust their behavior to a policy change such as delayed eligibility to social security than older respondents. Even leaving aside differences in preferences, we therefore predict more important changes in behavior for older respondents who do not have time to adjust their consumption over a long period of time. To see how this is translated in terms of the current model, consider Figure 6.1, presenting scatterplots of the predicted delay in retirement of our respondents by income (left panel) and age (right panel).[9] We see that at any given level of income and at any given age, the model predicts a substantial variation in behavior. In both cases, the average trends are correctly predicted.

Another way by which the respondents can adapt to the change in policy is by reducing their consumption.[10] The structural model used in this paper allows predictions on consumption and savings. It is therefore possible to predict changes in consumption that would be induced by the policy described above. Let us consider the immediate effect of the reform on consumption. To do so, let us consider the largest variation that the reform could induce in the model at the

9 One respondent with very low income was excluded from the graph. This respondent was predicted to delay retirement by almost two years.

10 Given the discrete nature of the model a consumption increase is possible. Given that a respondent can only retire once a year, he may decide to delay retirement and to increase consumption. This situation is never predicted by the simulations.

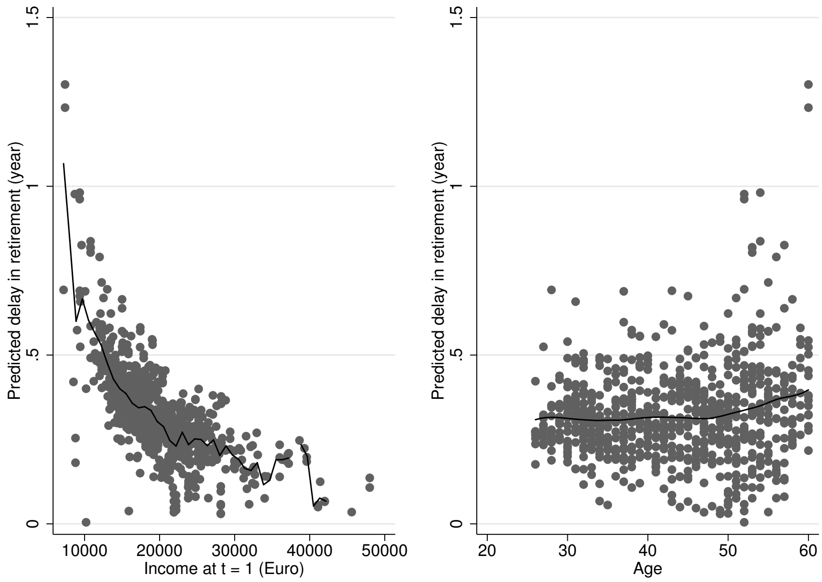


Figure 6.1: Predicted delay in planned retirement current income (left panel) and age (right panel), assuming that retirement is fixed at 65. The line is a non-parametric estimate of conditional expectations.

current time period and for a respondent planning to retire at 65. We compare the predicted consumption before and after the reform. This simulation is equivalent to the case of a respondent who did not forecast the change in policy and who learns of this change at the current time period. Given that most respondents seem to anticipate the reform to some extent (see Bissonnette and van Soest, 2011), and given that many of them may delay retirement, the difference in consumption path is likely to overestimate the observed effect and should be interpreted as an upper bound on predicted consumption change. Consider Figure 6.2, presenting scatterplots of the predicted variation by income (left panel) and by age (right panel). We hardly see a systematic effect of income on the immediate consumption decline, should retirement be fixed at 65. We also see that age is an important factor in the prediction. As we forecasted, older respondents would have to adapt their consumption more than younger respondents. However, remember that the current reform of AOW that should be implemented is gradual, and that a full delay in eligibility will only be in effect for respondents aged 55 or less. In this simulation, I supposed that everybody receives the same negative shock on their pension wealth, such that respondents aged over 55 have the same penalty as younger respondents. Therefore, the reaction to the policy reform should be smaller for older respondents than is simulated there. Note also that a lot of younger respondents would not change their consumption at all. This is due to the fact that they were prevented from borrowing from their pensions, and therefore still plan to enter

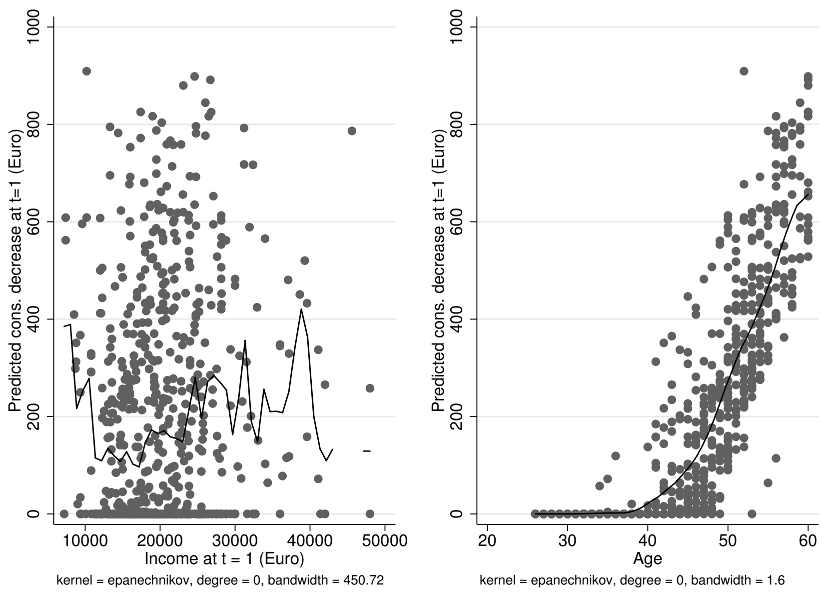


Figure 6.2: Predicted variation of consumption for current year by current income (left panel) and age (right panel), assuming that retirement is fixed at 65.

retirement with a value of accumulated assets equal to 0, as they did before.

To summarize these simulations, it seems that the model predicts that poorer respondents would tend to delay retirement more than higher earners should this reform pass, and to predict that a decrease of consumption is mostly caused by age. There is nevertheless a lot of heterogeneity in the predicted reaction to this reform.

6.6 DISCUSSION AND CONCLUSION

In this paper, I used simple questions where respondents are asked to state their preferences concerning hypothetical scenarios to estimate a life-cycle model of retirement and savings. I find that the estimated model yields a plausible estimation of parameters, generally in line with results from the literature, and showed that many of the problems I faced came from the small sample size or from unavailable data that could easily be elicited. Respondents seem to answer the questions in a way that is consistent over time, as seen from the importance of the random effects in the analysis. The scenarios I analyze are both easy to understand and are in line with the defined-benefit pension plans predominant in the Netherlands, which means that they may relate to them even if they do not have access to such retirement plans. Thus I could not find any obvious reason not to use stated preferences in the analysis of retirement and savings.

The next logical step for this research is to try to compare the predictions of this estimated model with actual behavior. As time goes by, we will be able to verify whether respondents who stated a higher preference for early retirement are also those retiring earlier, or will be able to use actual saving patterns to evaluate the adequacy of the model used. Failing to reconcile stated and revealed preferences would probably cast doubt on the usefulness of the stated-preference approach. Still, if what people report is what they plan or hope to achieve, or if it reflects their preferences at a given point in time, there may be some usefulness for the current analysis. A simple and straightforward application is to evaluate how a population may vote on some propositions. Understanding this is important if policy makers want to have the support of the population when proposing pension reforms, notably in states where votes and referendums are omnipresent. The state of California, for instance, is a good example.

For the sake of the discussion, let us assume that the information contained in the stated preference data is shown to be relevant for the study of retirement and savings. In this case, I will undoubtedly want to improve the simple model I presented here. Many of the shortcomings of this research are not due to the absence of an appropriate theory or to the high cost of eliciting the relevant data, but rather to the modest sample size available at the time of writing or to the lack of a required piece of information in the dataset. These problems are not a serious barrier to future research. Eliciting data of this type is not particularly expensive or challenging, mostly with access to panels such as the CentERpanel. There are a number of interesting and important factors that we may want to control for. The structural form dictates well how to augment the model, and, once the new value function is computed, the econometric model requires only minor adjustments, if any.

Future research topics may include better measurement of risk aversion or may rely on dynamically inconsistent preferences (e.g. Laibson, 1997). These last two examples would require cleverly designed future scenarios, both easy to understand for the respondents and complete enough to identify the model used. The fact that researchers can easily vary the assessed scenarios in an experimental manner should be a useful tool to actually identify these effects. Household structure and joint retirement are other interesting topics. I did not take into account, for instance, the number of children or dependent family members in each household. Again, the decision to exclude these characteristics from the model was mostly driven by the relatively small size of the sample available. We can nevertheless think of an extension to this model that would take these elements into account. In this case, we would probably want to allow the utility from consumption to vary according to the number of family members, but could presume that the value of leisure after retirement should not be affected (assuming that the children are no longer living with their parents, naturally). Finally, some health related variables, like self-assessed health, should probably be included in the leisure parameter. It would also be interesting to measure perceived spouse's health, and see how it affects retirement plans. Subjective survival probabilities could also be used in order to control for heterogeneity in frailty among the respondents. While these last variables were not available in the dataset, eliciting them does not pose a great challenge. Some of this information will be available in the 2011 yearly wave of the *Pension Barometer Survey*, allowing me (and hopefully other researchers) to improve the realism of the model.

6.7 ACKNOWLEDGEMENTS

I want to thank Rob Alessie, Charles Bellemare, Katie Carman, Marike Knoef, Marie-Louise Leroux, Pierre-Carl Michaud, Martin Salm, Arthur van Soest, and Hanka Voňková for valuable help and insightful comments. I am also grateful for funding by Netspar. Remaining mistakes are mine.

BIBLIOGRAPHY

Alessie, R., S. Hochguertel, and A. van Soest (2002): "Household Portfolio in the Netherlands," in *Household Portfolios*, ed. by L. Guiso, M. Haliassos, and T. Jappelli, pp. 341–388. MIT Press, Cambridge, MA.

Beggs, S., J. Cardell, and J. Hausman (1981): "Assessing the Potential Demand for Electric Cars," *Journal of Econometrics*, 17(1), 1–19.

Bertrand, M., and S. Mullainathan (2001): "Do People Mean What They Say? Implications for Subjective Survey Data," *The American Economic Review*, 91(2), 67–72.

Bissonnette, L., and A. van Soest (2010): "Retirement Expectations, Preferences, and Decisions," *Netspar Panel Paper*, 18, 1–98.

——— (2011): "The Future of Retirement and the Pension System: How the Public's Expectations Vary over Time and across Socio-Economic Groups," *CentER Working Paper Series*, (2011-065 5759).

Blau, D., and D. Gilleskie (2006): "Health Insurance and Retirement of Married Couples," *Journal of Applied Econometrics*, 21(7), 935–953.

Bovenberg, A. L., and R. H. J. M. Gradus (2008): "Dutch Policies Towards Ageing," *European View*, 7(2), 265–275.

Browning, M., and T. F. Crossley (2001): "The Life-Cycle Model of Consumption and Saving," *The Journal of Economic Perspectives*, 15(3), 3–22.

Browning, M., and A. Lusardi (1996): "Household Saving: Micro Theories and Micro Facts," *Journal of Economic literature*, 34(4), 1797–1855.

Capretta, J. (2007): "Global Aging and the Sustainability of Public Pension Systems: An Assessment of Reform Efforts in Twelve Developed Countries," *Center for Strategic & International Studies*.

De Bresser, J., and A. van Soest (2009): "Satisfaction with Pension Provisions in the Netherlands - A Panel Data Analysis," *Netspar Discussion Paper*, (DP 10/2009-034).

de Vos, K., and A. Kapteyn (2004): "Incentives and Exit Routes to Retirement in the Netherlands," in *Social Security Programs and Retirement Around the World*, ed. by J. Gruber, and D. A. Wise, vol. I, pp. 461–498. University of Chicago Press, Chicago.

Dominitz, J., and A. van Soest (2008): "Analysis of Survey Data," in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf, and L. E. Blume. Palgrave Macmillan, Basingstoke.

Euwals, R., A. Eymann, and A. Börsch-Supan (2004): "Who Determines Household Savings for Old Age? Evidence from Dutch Panel Data," *Journal of Economic Psychology*, 25, 195–211.

Fok, D., R. Paap, and A. van Dijk (2010): "A Rank-Ordered Logit Model with Unobserved Heterogeneity in Ranking Capabilities," *Journal of Applied Econometrics*.

Gan, L., M. Hurd, and D. McFadden (2005): "Individual Subjective Survival Curves," in *Analyses in Economics of Aging*, ed. by D. A. Wise, vol. I, pp. 377–411. University of Chicago Press, Chicago.

Gruber, J., and D. A. Wise (1999): *Social Security and Retirement around the World*. National Bureau of Economic Research, Inc, Chicago, IL.

——— (2004): *Social Security Programs and Retirement around the World: Micro-Estimation*. National Bureau of Economic Research, Inc, Chicago, IL.

Gustman, A. L., and T. L. Steinmeier (1986): "A Structural Retirement Model," *Econometrica*, 54(3), 555–584.

——— (2004): "What People Don't Know About Their Pensions and Social Security," in *Private Pensions and Public Policies*, ed. by W. G. Gale, J. B. Shoven, and M. J. Warshawsky, pp. 57–125. Brookings Institution, Washington, DC.

——— (2005): "Imperfect Knowledge of Social Security and Pensions," *Industrial Relations*, 44(2), 373–397.

Hurd, M. D. (1989): "Mortality Risk and Bequests," *Econometrica*, 57(4), 779–813.

——— (2009): "Subjective Probabilities in Household Surveys," *Annual Review of Economics*, 1(1), 543–564.

Knoef, M., R. Alessie, and A. Kalwij (2009): "Changes in the Income Distribution of the Dutch Elderly Between 1989 and 2020: A Microsimulation," *Netspar Discussion Paper*, (DP 09/2009-30).

Laibson, D. (1997): "Golden Egg and the Hyperbolic Discounting," *Quarterly Journal of Economics*, 62, 443–479.

Lusardi, A., and O. Mitchell (2007): "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education," *Business Economics*, 42(1), 35–44.

OECD (2011): *Pensions at a Glance 2011: Retirement-Income Systems in OECD and G20 Countries*. OECD Publishing, Paris.

Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65(4), 781–831.

Stock, J. H., and D. A. Wise (1990): "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 58(5), 1151–1180.

Van der Wiel, K. (2008): "Preparing for Policy Changes: Social Security Expectations and Pension Scheme Participation," *Netspar Discussion Paper*, (DP 07/2008-025).

——— (2009): "Have You Heard the News? : How Real-Life Expectations React to Publicity," *Netspar Discussion Paper*, (DP 06/2009-031).

Van Soest, A., A. Kapteyn, and J. Zissimopoulos (2006): "Using Stated Preferences Data to Analyze Preferences for Full and Partial Retirement," *IZA Discussion Paper*, (WP 2661).

Voňková, H., and A. van Soest (2009): "How Sensitive are Retirement Decisions to Financial Incentives: A Stated Preference Analysis," *Netspar Discussion Paper*, (DP 10/2009-036).

Yaari, M. (1965): "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *The Review of Economic Studies*, 32(2), 137–150.

6.A Assets and debts definitions

I list here the list of assets and debt used in the aggregated definition. This list is the same as presented in some work I co-authored (Bissonnette and van Soest, 2010). The value of accumulated asset used in computation is the difference between wealth and debt at the beginning of the year. The elements included are as follows:
Wealth is the sum of the total amounts of the following elements:

1. Checking account
2. Employer-sponsored saving plan
3. Savings accounts
4. Deposit books
5. Savings certificates
6. Single-premium insurance policies
7. Endowment insurance policies
8. Growth funds
9. Bonds
10. Shares
11. All kinds of options
12. Real estate not used for housing
13. Money lent to family
14. Investment not mentioned before
15. Stocks for substantial holding
16. Business equity

The debt is the sum of the total amounts of these elements:

1. Private loans
2. Extended lines of credit
3. Debt with mail-order firms
4. Hire-purchase contracts
5. Loans from family and friends
6. Study loans
7. Credit card debts
8. Loans not mentioned before

6.B ON THE DUTCH PENSION SYSTEM

The description of the Dutch pension system contained in this part is based on more detailed articles written by de Vos and Kapteyn (2004) and Bovenberg and Gradus (2008). Other useful sources of information are the yearly publication by the OECD titled *Pension at Glance* (e.g. OECD, 2011) and the report written by Capretta (2007). I refer readers interested in additional details to these publications. Before going further, I want to stress that the data used in this paper were elicited between 2006 and 2008. Hence, some more recent changes discussed by Bovenberg and Gradus may not have been in place when the respondents answered the survey.

Let us think about the Dutch pension system in terms of three pillars. The first pillar, the AOW, is a minimum pension provided by the government, and is a form of social security. This public pension is linked to the minimum wage. For instance, singles receive 70% of the minimum wage after the age of 65 and an individual in a two-person household receives 50% of this amount. Then, as a second pillar, a large fraction of the labor force (about 90% according to Bovenberg and Gradus) are entitled to an occupational pension. This large proportion is explained by the fact that participation to the pension scheme is mandatory if an employer offers it. According to de Vos and Kapteyn: "until recently, more than 99 percent of the pension schemes were of the defined-benefit type, most of them being defined on the basis of final pay". Bovenberg and Gradus report that career-average schemes, however, are becoming more common. According the OECD, the first and second pillar are often integrated, so that workers are offered a 70% replacement rate if they retire at age 65 and had a stable career. Finally, third pillar pension, individual provisions, is relatively small in the Netherlands.

Another important element in the Dutch system is the large prevalence of early retirement, due to a very generous pay-as-you-go option (called VUT) that were first set in place with the intention to curb unemployment by increasing the number of retiree. Given the high costs associated with this program, it was decided in 2005 that VUT would be abolished. Nevertheless, according to Capretta (2007): "With over two-thirds of men (and over four-fifths of women) exiting the workforce by age 60, early retirement is still the norm." It should also be noted that reaching age 65 is a legitimate reason for dismissal and that it is not possible to delay payment of AOW beyond this age. In effect, 65 is the mandatory retirement age for most occupations. Hence, whereas Dutch in the labor force may plan to retire early, they may not perceive that staying in the labor force past 65 is an option for them, or at least can consider it as riskier than usually acknowledged. A brief analysis on this point is presented by van Soest et al. (2006). Respondents in their sample were asked the earliest and latest age at which they could retire according to their employer's pension plans. They report that the earliest age varies from 55 to 65 (median of 62, mean of 61.7) and that latest retirement was concentrated at age 65. Another important aspect is that the most frequently used answer for early retirement is also 65, implying that a large fraction of respondents do not think that they have any flexibility regarding their retirement age.